NBER WORKING PAPER SERIES

NONLINEAR PRICING IN VILLAGE ECONOMIES

Orazio Attanasio
Elena Pastorino

Working Paper 21718
http://www.nber.org/papers/w21718

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2015, Revised July 2018

We are grateful to V.V. Chari, Aureo de Paola, Tom Holmes, Bruno Jullien, Toru Kitagawa, Aprajit Mahajan, Costas Meghir, Matt Mitchell, Mark Rosenzweig, Holger Sieg, Kjetil Storesletten, Chris Taber, Joel Waldfogel, and Frank Wolak for helpful suggestions, and several seminar audiences for comments. Eugenia Gonzalez Aguado, Sam Bailey, Ross Batzer, Javier Brugues- Rodriguez, Sanghmitra Gautam, and Pedro Veloso have provided superlative research assistance. We especially thank Kai Ding, Keyvan Eslami, and Rishabh Kirpalani for many useful comments and Joan Gieseke for her invaluable editorial assistance. The paper is a greatly revised version of Attanasio and Pastorino (2010). The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis, the Federal Reserve System, or the National Bureau of Economic Research.

Nonlinear Pricing in Village Economies
Orazio Attanasio and Elena Pastorino
NBER Working Paper No. 21718
November 2015, Revised July 2018
JEL No. D42,D43,D82,I38,O12,O13,O22

ABSTRACT

This paper analyzes and estimates the impact of quantity discounts for basic staples in rural Mexico. We propose a model of price discrimination that nests those of Maskin and Riley (1984) and Jullien (2000), in which consumers differ in their tastes and, due to subsistence constraints, in their ability to pay for a good. We show that, under mild conditions, a model in which consumers face heterogeneous subsistence or budget constraints is equivalent to one in which consumers have access to heterogeneous outside options. We then rely on known results (Jullien (2000)) to characterize the equilibrium price schedule. We analyze the effects of nonlinear pricing on market participation, as well as the impact of a market-wide income transfer on the price schedule, when consumers are differentially budget constrained. Such a transfer, which is a common policy in developing countries, stimulates consumption by increasing households' ability to pay but also typically leads to an increase in the intensity of price discrimination. We prove that the structural parameters of the model are identified from data on prices and quantities in a given market. The intuition behind this identification result is that the equilibrium price schedule in a market is a function of the distribution of consumers' willingness and ability to pay, which, in turn, is related to the distribution of quantities. We estimate the parameters of the model using data from municipalities and localities in Mexico on three commodities, rice, kidney beans, and sugar, which are consumed by most households and are part of the evaluation component of the well-known conditional cash transfer program Progresa. The model fits the data remarkably well. Interestingly, we find that nonlinear pricing is beneficial to a large number of households, especially those who consume small quantities, relative to linear pricing. Available evidence indicates that the program had no effect on average prices. We show, however, that the program has affected the slope (in quantity) of the price schedule, which has become steeper as implied by our model, and has thus lead to an increase in the degree of price discrimination. We also show that, empirically, accounting for the impact of the hazard function of the distribution of quantities on prices, as consistent with our model, explains a large fraction of the shift in the price schedule induced by the program.

Orazio Attanasio
Department of Economics
University College London
Gower Street
London WC1E 6BT
UNITED KINGDOM
and NBER
o.attanasio@ucl.ac.uk

Elena Pastorino
Stanford University
Hoover Institution and Economics Department
579 Serra mall
94305
Stanford University
Palo Alto, CA 94305
and the Federal Reserve Bank of Minneapolis
epastori@stanford.edu

A supplementary appendix is available at http://www.nber.org/data-appendix/w21718

1 Introduction

Quantity discounts, in the form of unit prices declining with quantity, appear to be pervasive in developing countries. McIntosh (2003), for instance, documents differences in the price of drinking water paid by poor households in the Philippines, whereas Fabricant et al. (1999) and Pannarunothai and Mills (1997) document similar differences in the price of health care and services in Thailand and Sierra Leone. Attanasio and Frayne (2006) show evidence that consumers purchasing basic staples in Colombian villages face price schedules rather than linear prices: within a village, relatively richer households buy larger quantities and pay substantially lower unit prices for homogeneous commodities. Similar patterns arise in rural Mexico, as we document here. This evidence is commonly interpreted as symptomatic of poor households paying more than rich ones for the same goods they purchase. Since the poor are close to subsistence in developing countries, by this argument, nonlinear prices are usually considered to have undesirable distributional implications.

This view is consistent with the implications of the nonlinear pricing model of Maskin and Riley (1984), henceforth the *standard model*, in which consumers differ in their marginal willingness to pay for a good. This model explains quantity discounts as arising from a seller's incentive to screen consumers by their preferences through the offer of multiple price and quantity combinations. One of its key insights is that the ability of a seller to discriminate across consumers implies not only that the consumption of (nearly) all consumers is depressed relative to first best but also, crucially, that consumption distortions tend to be more severe for purchasers of smaller quantities, typically poorer consumers in developing countries.

The standard model, however, assumes consumers differ only in their tastes, are unconstrained in their ability to pay, and have access to similar alternatives to purchasing from a particular seller or in a given market. This framework thus naturally accounts for the dispersion in the unit prices of goods that absorb a small fraction of consumers' incomes, in settings in which consumers have access to similar outside consumption opportunities.[1] As such, the standard model abstracts from key features of food markets in developing countries, in which households typically face subsistence constraints on the consumption of basic staples, spend a large fraction of their income on food, and have access to alternative consumption possibilities, through self-production or highly subsidized government stores, that are uncommon in developed countries.

To rationalize the occurrence of quantity discounts in settings that are typical of developing countries, we propose a model of price discrimination that explicitly formalizes households' subsistence constraints and allows households to differ in both their marginal willingness to pay and their absolute ability to pay for a good. The model also incorporates a rich set of alternatives to purchasing in a particular market that differ across consumers. The model we propose is structurally identified from information just on the distribution of prices and quantities in a market. We estimate the model on data from a sample of villages in rural Mexico collected to evaluate the conditional cash transfer program Progresa and show that it successfully fits the observed differences in prices and quantities consumed within and across villages. We find that, contrary to common intuition, nonlinear pricing is beneficial to a large number of households, especially

[1] Formally, consumers are assumed to be able to pay more than their reservation prices for a good. See Che and Gale (2000).

those who consume less, relative to linear pricing. We also characterize and estimate the effect of income (cash) transfers in this framework and find that in the villages we study, the Progresa transfers stimulated consumption but also induced an increase in the intensity of price discrimination, which affected negatively both beneficiaries of the transfers as well as non-eligible households.

This paper makes four important contributions. First, from a theoretical point of view, we show that when facing subsistence constraints, consumers can be formally thought of as facing an additional budget constraint on the expenditure on a seller's good. In the language of the literature on auctions and nonlinear pricing, consumers are *budget constrained*, and their constraints depend on their preferences and incomes. This class of models has been considered in general to be intractable. We show that even when consumers differ in both their tastes and budgets, a model with budget-constrained consumers maps into the class of nonlinear pricing models with so-called countervailing incentives, in particular the one of Jullien (2000), in which consumers have heterogeneous outside options. By relying on this formal equivalence between models, we can exploit existing results to characterize optimal nonlinear pricing. In such an environment, nonlinear pricing has more desirable welfare properties than those implied by the standard model, as it leads to higher levels of consumption, even above first best, lower marginal prices, and for sufficiently comparable outside options across the two models, higher consumer surplus. Intuitively, since the existence of budget constraints implies that there is a maximum price that a consumer can pay, budget constraints effectively limit sellers' ability to extract consumer surplus. In these instances, nonlinear pricing may be preferable to linear pricing, especially because it may lead to a higher degree of market participation.

Second, we show that the model's primitives and structural parameters are identified just from information on the distribution of prices and quantities in one market. The intuition behind this result is relatively simple. In the model, a price schedule is set by a seller to discriminate among consumers with different tastes and budgets, and therefore it depends on the distribution of consumers' characteristics. Since this latter distribution is reflected in that of quantities, the observed distribution of quantities can be used to pin down the determinants of a price schedule. The estimation approach we propose crucially relies on the seller's optimality conditions, which imply that the difference between prices and marginal costs depends on the difference between the cumulative multiplier associated with consumers' participation constraints and the distribution function of consumers' preferences. The approach we propose, relying on this relationship, allows us to discriminate between different versions of the model, including the standard model that our model nests, and therefore to identify the distinct welfare implications of different types of price schedules.

Third, we estimate the model for three commodities in a large number of villages in rural Mexico. We use data from the high-quality surveys collected for the evaluation of the conditional cash transfer program Progresa, which have been extensively used. Although we can achieve nonparametric identification, to improve efficiency and be able to deal with the small number of observations in several villages, we choose specific yet flexible functional forms for the relevant objects. We show that our model fits remarkably well the data we use and is robust to key parameterizations. In particular, the estimates of the model's primitives satisfy the model's restrictions on the monotonicity of the hazard rate of the distribution of consumers' unobserved marginal willingness to pay, as well as the inverse relationship between marginal utility and quantity consumed, without being imposed. We are also able to test different versions of the model and show that the standard model is rejected in all villages.

Fourth, although the effect of conditional cash transfer programs including Progresa on *average* prices has been studied before and often detected to be negligible or absent, we show that the program has induced significant shifts in *price schedules* that might have had important distributional implications. Furthermore, using an approximate reduced-form version of the optimality conditions for sellers' and consumers' behavior, we show that our model is able to explain a large fraction of the shift in price schedules we document.

The intuition for some of our theoretical results is simple and deserves a mention, in particular, the impact of nonlinear pricing on market participation and of income transfers on prices. As for market participation, we show that nonlinear pricing can be efficiency enhancing relative to linear pricing by leading to greater market participation, when consumers are differentially constrained in their access to a market. Specifically, by allowing a seller to tailor prices and quantities to consumers' willingness (and ability) to pay, nonlinear pricing enables a seller to trade at a profit with consumers with more stringent subsistence constraints—typically poorer consumers purchasing smaller quantities—or with access to especially attractive outside options. Such consumers would be excluded from the market under linear pricing and so are better off under nonlinear pricing. The reason is as follows. To induce such consumers to participate, a seller needs to offer a low enough marginal price. Since the marginal price is constant and equals the unit price under linear pricing, such a low linear price would lower profits from all consumers for the benefit of including just a few more. Thus, including these consumers would not typically be profitable at linear prices.

In terms of the impact of income (cash) transfers on prices, we show that these policies not only encourage consumption but also provide an incentive for sellers to take advantage of consumers' greater ability to pay. In particular, targeted cash transfers that are more generous for poorer households, who tend to purchase smaller quantities, lead not just to a greater demand for a seller's good but also to higher prices for at least some consumers. Depending on the distribution of tastes and outside options across consumers, transfers can give rise to increases in the unit prices of low quantities but decreases in the unit prices of high quantities. Although the average impact on unit prices is therefore ambiguous, we prove that overall transfers lead to an increase in the degree of price discrimination, thereby exacerbating some of the consumption distortions associated with nonlinear pricing.

Our empirical results are consistent with these intuitions. Our estimates imply that sellers have substantial market power in the villages in our data and exercise it by price discriminating across consumers through distortionary quantity discounts. However, the pattern of consumption distortions we detect is mostly *opposite* to that implied by the standard model. In particular, in most villages, the consumption of households with intermediate to large taste for a good is the most compressed relative to first best, whereas a sizable fraction of consumers of low to intermediate quantities, with low to intermediate tastes, do not suffer large distortions. Under the standard model, instead, these latter consumers would suffer the largest distortions. Interestingly, several consumers of the smallest quantities consume quantities *above* first best rather than, as often argued, below it—in the standard model, as mentioned, consumption is below first best for all consumers.

When comparing observed nonlinear pricing to a counterfactual scenario in which sellers cannot price discriminate, we find that nonlinear pricing leads to greater consumer and social surplus for most consumers with low to intermediate types, especially consumers of the smallest quantities. A large fraction of such consumers would be excluded under linear pricing and thus benefit from nonlinear pricing. On the contrary, consumers of large quantities tend to be better off under linear pricing.

Several studies have analyzed the effect of transfers on the price of commodities and have found no impact. Hoddinott et al. (2000), for instance, conclude that "there is no evidence that Progresa communities paid higher food prices than similar control communities" (p. 33). Similarly, Angelucci and De Giorgi (2009), when assessing the impact of Progresa on the consumption of non-eligible households, consider the possibility that their results are mediated by changes in local prices but dismiss this possibility based on their empirical analysis. Thus, the consensus seems to be that Progresa did *not* have noticeable effects on local prices. Our finding that cash transfers implemented by Progresa have had a significant impact on prices (and in particular on price schedules) is therefore novel, as far as we know. Importantly, we show that the effects of the program on the price schedule that we document can be explained by our model.[2]

The reason for the difference between our results and those in the existing literature is as follows. Most of the studies that have looked at the effect of cash transfers on prices focus on *average* changes in unit prices associated with the introduction of cash transfers but fail to account for the nonlinearity of unit prices when assessing the impact of transfers on prices. Our model implies that income transfers to consumers affect equilibrium prices, as sellers adjust their price schedules in response to consumers' greater incomes. In line with this prediction, we find that the schedule of unit prices becomes significantly *steeper* after transfers are introduced, with unit prices increasing at low quantities but decreasing at high quantities. Since the resulting *average* impact of cash transfers on unit prices is much less pronounced—in fact, insignificant according to our estimates—we also show that ignoring the variability of unit prices with quantity leads to a much smaller estimate of the price effect of transfers, as consistent with the results in the literature. When, instead, the dependence of unit prices on quantity is explicitly taken into account, the price effect is substantial: by increasing consumers' ability to pay, cash transfers also effectively increase a seller's market power and the intensity of price discrimination. Such a change in equilibrium prices affects not just households who are beneficiaries of the program but also non-eligible households. Cash transfers can then lead to much lower consumer surplus gains than typically inferred.

As for the rest of the paper, in Section 2, we discuss our sample of rural villages included in the evaluation of Progresa. In Section 3, we present our model, characterize optimal nonlinear pricing, and analyze its implications for consumption, market participation, and the impact of income transfers. In Section 4, we prove that our model is identified and propose an estimation strategy. In Section 5, we present our estimates, assess their welfare implications, and evaluate the effect of the Progresa transfer on prices. Finally, Section 6 concludes the paper.

[2] Interestingly, Cunha et al. (2017), who analyze a different program that included cash and in-kind transfers, find that the latter significantly reduced prices. Once again, this pattern is consistent with our model, as discussed in Section 3.

2 Quantity Discounts: The Case of Mexico

Here we provide a description of our data from the Progresa program and evidence of quantity discounts.

Data: Background and Description. The dataset we use was collected to evaluate the impact of the conditional cash transfer program called Progresa, which was started in 1997 under the Zedillo administration in Mexico. The program consists of cash transfers to eligible families with children, conditional on behavior such as class attendance by school-aged children and regular visits by young children to health centers, as well as attendance of education sessions on nutrition and health by mothers. Among the villages targeted by the first wave of the program expansion between 1998 and 2000, about 70% of households were deemed eligible. Eligibility was determined on the basis of a survey that collected information on a set of poverty indicators considered difficult to manipulate, such as the material of the roof or floor of a household's home.[3]

In this first phase targeted to marginalized rural communities, the Progresa grant consisted of two components. The first one was targeted to families with children less than 6 years old and was conditioned on children being brought to health centers with some regularity. The second component was targeted to families with children between ages 9 and 16 and was conditioned on regular school attendance. Although the program administration was relatively strict in enforcing these conditionalities, they were not very binding for several households—for instance, households with primary school aged children whose school attendance is very high. For eligible households, the grant was substantial. On average, grants amounted to 25% of their income and consumption, therefore constituting a large fraction of it.

Since the first rollout of the program was targeted to about 20,000 marginalized localities and would take about two years to be implemented, the program's administration and the government decided to use it for evaluation purposes, by randomizing the timing of part of the rollout. In particular, in 1997 the program selected 506 *localities* in 7 states, each belonging to one of 191 larger administrative units, *municipalities*, to be included in the evaluation sample. Each municipality is composed of several localities, not all of which were included in the evaluation survey. Of these localities, 320 were randomly chosen and assigned to early treatment in that the program started there in the middle of 1998; the remaining 186 were assigned to the end of the rollout phase, so the program started there in December 1999. Households in these localities were followed for several periods. We use the waves of October 1998, May and November 1999, and November 2000 and 2003. Some waves, such as those collected in 1997 and April 2000, could not be used because they do not contain information on household expenditure, which we rely on.

The evaluation data have been used extensively in recent years and are remarkable for at least three reasons. First, the randomized rollout of the program in a subset of the villages—at least for the first waves—introduced substantial exogenous variation in the resources available to some households, which we exploit to examine key implications of the model we propose, in particular about the impact of cash transfers on prices. Second, the data provide a census of 506 villages in that all households in the relevant localities are

[3] A first registration wave in 1997 was complemented by some further registrations in early 1998, the so-called *densificados*, as the program administration assessed eligible households to be too few, which led to a slight modification of eligibility rules. We consider these added families as eligible.

surveyed, therefore allowing us to estimate the entire distribution of quantities and prices in each village, at least for commodities whose purchase is common. Third, the data are very rich and exhaustive. The consumption and expenditure module in the survey contains key information for the purpose of our paper. Each household is interviewed and asked to report not only the quantity consumed of each of 36 food commodities during the week preceding the interview, but also the quantity purchased and its monetary value. The data also contain information about quantities consumed and not purchased, for instance, acquired through self-production or received as a gift or a payment in kind. The food items recorded include fruits and vegetables, grains and pulses, and meat and other animal products, and are supposed to be exhaustive of the foods consumed by households. In what follows, we focus on commodities that are relatively homogeneous in terms of quality, and are purchased and consumed by most households, as explained below.

Since the survey contains information on the quantities purchased of each recorded item as well as on total household expenditure on each item, it is possible to compute *unit values*, as measured by the ratio of expenditure to quantity purchased. From now on, we refer to unit values as *prices*. Attanasio et al. (2013) discuss some of the measurement issues associated with the construction of unit values, ranging from measurement error to the heterogeneous quality of goods and to the nonlinearity in quantity we model here. However, they find that average and median unit values well approximate local prices, which are available for some commodities in the locality survey and are collected from local stores. They also find that unit values closely match data from national sources on prices.

Progresa was targeted to marginalized communities, where marginalization was defined by an index used by the Mexican government to target social programs. However, they were not the most marginalized communities in the country. The exclusion of the poorest communities (targeted by a different program, studied, for instance, by Cunha et al. (2017)) was justified by the fact that to comply with the Progresa conditions, eligible households had to have access to certain public services and infrastructure, such as schools and health centers. The communities included in the evaluation survey are small—the average number of households in a locality is just over 50—and remote. Households living in these villages are poor: on average, for instance, food accounts for nearly 70% of household budgets. However, within villages, the level of poverty exhibits substantial variation. These differences are captured by a variety of indicators and reflected in the fact that not every household within a village was eligible for Progresa: on average, about 78% of the households of the villages in the evaluation survey were. In addition to variation in poverty within villages, our sample is also characterized by variation in the level of poverty *across* villages. This heterogeneity is reflected, for instance, in the variability of the rate of eligible households across villages, which is obviously related to differences in the distribution of consumption. Such heterogeneity should thus account for differences in price schedules across villages. Furthermore, differences in the proportion of eligible households across villages are likely to affect how the Progresa transfer has modified the distribution of quantities consumed and, according to our model, the shape of a village price schedule. We will investigate these relationships in our empirical analysis below.

To perform the empirical exercise we propose, we develop a model that relates the shape of the price schedule to the distribution of quantities in a given market. Since we specialize the model to focus on the behavior of a single seller in our empirical analysis, ideally, one would like to consider a relatively isolated market with one seller or a small number of them. Although the definition of a market as a locality would be natural, using the locality as our definition of "market" would result in few observed transactions in several cases, given the size of localities and the number of transactions we observe. Moreover, despite some localities being quite isolated, from an administrative point of view, they all belong to a municipality and are often connected in several ways. For instance, it is not unusual for some households to shop for certain items in a locality within the municipality of residence but different from the locality where they live. For these reasons, in the main text we focus primarily on villages defined as municipalities. However, we estimate our model on both village markets defined as municipalities and on village markets defined as localities and obtain fairly similar results across these two definitions of villages, as discussed in Section 4.

Evidence of Quantity Discounts. The model of nonlinear pricing we propose gives rise to quantity discounts. As mentioned in the introduction, this type of discounts is common in several markets in developing countries. Attanasio and Frayne (2006), for instance, estimate the supply schedule for several basic food staples, including rice, carrots, and beans in Colombian villages, and document substantial discounts for large volumes. Specifically, they find that the elasticity of the unit price of rice to the quantity bought is as large as -0.11 in their preferred specification. They estimate even larger quantity discounts for other specifications and for commodities such as carrots or beans. In our empirical exercise, we use the Mexican data described above. To motivate the model we develop and estimate using these data, we start by documenting the existence of discount patterns in Mexico similar to those observed in Colombia.

Table 1: Price Schedule

	Rice	Kidney Beans	Sugar	Tomatoes	Tortillas
Constant	1.858	2.426	1.795	2.256	2.080
	(0.007)	(0.008)	(0.006)	(0.010)	(0.046)
Ln(quantity)	-0.247	-0.145	-0.172	-0.320	-0.452
	(0.013)	(0.011)	(0.017)	(0.019)	(0.020)
Observations	$62,368$	$82,024$	$91,782$	$108,345$	$38,599$
R^2	0.139	0.094	0.116	0.331	0.353

Note: wave fixed effects are included; standard errors are clustered at the locality level.

For a sense of quantity discounts and their importance in our data, we relate observed prices to quantities purchased for several commodities in each village. Here we use the waves of October 1998, May and November 1999, and November 2000 and 2003 from all localities. Table 1 contains estimates of this relationship between (real) log unit prices and log quantities for the most common commodities, namely, rice, kidney beans, sugar, tomatoes, and tortillas. The different number of observations in each row reflects the different number of purchases we observe in our sample. The elasticity of prices to quantities we observe is largest (in absolute value) for tortillas (-0.45), tomatoes (-0.32), and rice (-0.25). However, it is also sizable for the other two commodities: -0.14 for kidney beans and -0.17 for sugar. For all these commodities, this elasticity is statistically different from zero. The estimates we obtain do not vary appreciably when we restrict attention to the somewhat smaller set of municipalities and localities part of our estimation sample, discussed below. As explained in Section 4, the sample selection criteria we impose when estimating the model entail only a small loss of villages.

Some studies have argued that the dispersion in prices observed in a market for a same commodity might reflect differences in quality. Deaton (1989), for instance, argues that this might be the case for rice in Thailand. Here we focus on commodities for which the quality homogeneity assumption does not seem unreasonable in our context for two reasons. First, even in the case of rice, conversations with program officials and others familiar with the reality of rural Mexico confirm that this is the case. Second, we do not consider commodities such as meat, for which quality is likely to be heterogeneous. In any case, heterogeneity in the quality of goods would likely give rise to an *upward-sloping* price schedule, contrary to what we observe; see also Crawford et al. (2003). The evidence in Table 1 therefore seems to indicate the presence of important quantity discounts for the commodities considered.

For the rest of our empirical analysis, we focus on three commodities—rice, kidney beans, and sugar—that conform to the assumptions we maintain in the theoretical model. Specifically, we consider commodities that are of homogeneous quality, so as to minimize the possibility that price differences reflect any heterogeneity in this unmodeled dimension. We also focus on commodities that are commonly consumed and purchased by the households in our sample. Of the commodities in Table 1, we exclude tortillas because a large number of households buy corn or wheat to make them and because price differences are known to reflect quality differentials. Tomatoes are also excluded because of potential quality differentials affecting their prices that we cannot detect in our data. The commodities we choose are not only widely consumed but also storable, so observed zero purchases do not reflect exclusion from the market but could simply be due to the timing of the Progresa interview. Hence, the assumption of full market participation we formulate in our empirical analysis is not implausible. Across localities, the median fraction of households consuming rice in the week preceding the interview is 58%, while the corresponding fraction is 85% for kidney beans and sugar. Virtually all of these households (97% for rice, 91% for kidney beans, and 98% for sugar) purchase these commodities rather than producing them or receiving them as a gift or an in-kind payment.

The model we present in the next section considers a seller facing a heterogeneous population of consumers with any degree of market power, ranging from monopoly to oligopoly to no power (perfect competition). In the empirical analysis, however, we specialize our model to consider the behavior of only one seller in each market. This assumption is for consistency with the data, which show that markets defined as either municipalities or localities are highly concentrated with very few stores. Specifically, in the 506 localities in our dataset, the median number of stores is 1 or 2 depending on the Progresa wave. As for municipalities, the mean and median number of stores are higher, as some government stores and other very heterogeneous types of sellers, such as periodic open air and itinerant street markets, might be present. These sellers, however, can be considered as having a very different cost structure, and their possible presence in the markets for the commodities we study can be interpreted as a degree of competition that is incorporated into households' outside options in our model. Moreover, even at the level of the municipality, the number of grocery stores that might sell the goods we consider is very small: the median is zero and the mean is 0.68. Hence, the supply of the goods we focus on is indeed highly concentrated in each market, so the restriction to one seller per market in our empirical analysis is not a strong one.

3 Models of Price Discrimination

As just reviewed, for many commodities in rural Mexico, prices per unit decline with the quantity purchased. A simple model consistent with this feature of the data is the *standard model* of price discrimination of Maskin and Riley (1984), in which quantity discounts emerge when a seller screens consumers by their marginal willingness to pay according to the quantities they purchase. This model, however, can be too restrictive for the markets we study, since it assumes that consumers have the same reservation utilities and abstracts from consumers' budget or subsistence constraints. To capture the value of richer consumption possibilities alternative to purchasing from a particular seller or in a particular market, we build on the model of Jullien (2000), which assumes consumers differ not just in their taste for a good but also in their reservation utility. Suitable interpretations of consumers' reservation utility can then accommodate different settings of interest.

A particularly relevant case arises when consumers face subsistence constraints in consumption, which give rise to a *budget constraint* on the expenditure on a seller's good. As discussed in Che and Gale (2000), models with this type of budget constraint are considered to be intractable in general. The optimal pricing schedule is only known for special cases, when, for instance, utility is linear in consumption (see Che and Gale (2000)) or the budget is identical across consumers, so that a seller does not have an incentive to discriminate across all consumers (see Thomas (2002)). In what follows, we establish that a model with heterogeneous budget constraints is equivalent to a model with heterogeneous reservation utilities under simple conditions. This equivalence allows us to adapt and extend the results in Jullien (2000) to a model with budget-constrained consumers and characterize nonlinear pricing in the presence of these constraints.

The model we propose captures a variety of situations that are relevant to our application. First, consumers in our data have access to a wide range of outside options: households in a village may purchase a good from sellers in other villages or in government-regulated *Diconsa* stores; they may have the ability to produce a good as an alternative to purchasing it; or they may receive a good from relatives, friends, or the government as a transfer. As the desirability or feasibility of these alternative consumption possibilities may differ across consumers, so does consumers' reservation utility. Second, although we focus on the problem of a single seller, by interpreting a consumer's reservation utility as the utility obtained when purchasing from *other* sellers, the model can account for varying degrees of seller market power and so different market structures, ranging from monopoly to oligopoly to near perfect competition.[4] Finally, as consumers typically have preferences for multiple goods, we incorporate the possibility of consumers' substitution across them and allow subsistence constraints to affect the consumption of any good.

[4]Near competition obtains when the reservation utility is arbitrarily close to first-best utility. The problem of a seller we consider can be alternatively interpreted as the best-response problem of a price-discriminating oligopolist competing to exclusively serve any given consumer in a village. We establish this result in the Supplementary Appendix.

As is common in the nonlinear pricing literature, our framework implicitly excludes the possibility of collusion among consumers, for instance, through resale. Anecdotal evidence we obtained from program officers and surveyors indicates that resale does not occur in our context. A natural question is why consumers do not form coalitions, buy in bulk, and resell quantities among themselves at linear prices. A possible answer is that our context is that of small, isolated, and geographically dispersed communities in rural Mexico. Thus, it might be difficult for consumers to engage in the type of agreements that would sustain resale.[5]

3.1 A Model with Heterogeneous Outside Options

Consider a market (village) in which consumers (households) and a seller exchange a quantity $q \geq 0$ of a good for a monetary transfer t. Consumers' preferences depend on a taste attribute, θ, continuously distributed with support $[\underline{\theta}, \overline{\theta}]$, $\underline{\theta} > 0$, cumulative distribution function $F(\theta)$, and probability density function $f(\theta)$, positive for $\theta \in (\underline{\theta}, \overline{\theta})$. We refer to this attribute as *marginal willingness to pay*. We assume the seller knows the distribution of θ but does not observe its value for a given consumer or, alternatively, prices contingent on consumers' characteristics are not legally permitted or enforceable. Thus, a seller must post a single price schedule for all consumers, although it can entail different unit prices for different quantities.[6]

Faced with a price schedule, a consumer decides whether to purchase and, if so, the quantity q to buy. Upon trade, a consumer of type θ obtains utility $v(\theta, q) - t$, with $v(\cdot, \cdot)$ positive and twice continuously differentiable, $v_\theta(\theta, q) > 0$, $v_q(\theta, q) > 0$, and $v_{qq}(\theta, q) \leq 0$. We assume, as standard, that $v_{\theta q}(\theta, q) > 0$ for $q > 0$ so that consumers can be ordered by their marginal utility from the good. Denote by $c(\cdot)$ the seller's cost function, which is weakly increasing and twice continuously differentiable, and by $c(Q)$ the cost of producing the total quantity of the good provided to the market, Q. For simplicity of exposition, here we maintain that the cost function $c(\cdot)$ is additively separable across consumers; we relax this assumption in the empirical analysis.[7] We denote by $s(\theta, q) = v(\theta, q) - c(q)$ the *social surplus* from quantity q and maintain that $s_q(\theta, \cdot)/v_{\theta q}(\theta, \cdot)$ decreases with q. This assumption ensures that the seller's problem admits a unique solution and that first-order conditions are necessary and sufficient to characterize it: it plays the same role as the assumptions that $s(\theta, \cdot)$ is concave in q and $v_\theta(\theta, \cdot)$ is convex in q in the standard model. We define the *first-best* quantity, $q_{FB}(\theta)$, as the one that maximizes social surplus for a consumer of type θ.

Let $\overline{u}(\theta)$ be a consumer's *reservation utility* when not purchasing from the seller, which is assumed to be absolutely continuous and, unlike in the standard model, can vary across consumers. We normalize the seller's reservation profit to zero. A consumer of type θ *participates* when the consumer purchases a single quantity with probability one—the restriction to deterministic contracts is without loss. We focus on situations in which all consumers trade, so $q = 0$ is interpreted as the limit when the contracted quantity

[5] Conceptually, such a situation can be translated into a simple assumption on the existence of imperfections in contracting between consumers analogous to those between sellers and consumers usually maintained in models of nonlinear pricing. With enforcement, coordination, or transaction costs, such as commuting costs, not even a coalition of all consumers could achieve higher utility for any member than the utility a member obtains by trading with a price-discriminating seller. See Appendix A.

[6] We rely on results from the mechanism design literature with private information. A standard result, the *taxation principle*, is that an economy with observable types in which a seller is restricted to nonlinear prices, referred to as "tariffs," is equivalent to an economy with unobservable types and no restrictions on the space of contracts a seller can offer. See Segal and Tadelis (2005).

[7] One instance of nonseparability occurs when the cost function is subadditive. Given the compact support of consumer types, with a suitable lower bound on $c'(Q)$, we can ensure that the seller's problem is concave even in this case.

becomes small. Observe, however, that in the presence of consumer exclusion, the equilibrium contract for types who participate would be the same as the one we characterize below.

By the revelation principle, a contract between consumers and seller can be summarized by a menu $\{t(\theta), q(\theta)\}$ such that the best choice within the menu for a consumer of type θ is the quantity $q(\theta)$ for the price $t(\theta)$; that is, the menu is *incentive compatible*. Let $u(\theta) = v(\theta, q(\theta)) - t(\theta)$ denote the utility of a consumer of type θ when purchasing from the seller under the incentive compatible menu $\{t(\theta), q(\theta)\}$. The seller's optimal menu maximizes expected profits subject to consumers' incentive compatibility and participation constraints, that is,

$$\text{(IR problem)} \qquad \max_{\{t(\theta),q(\theta)\}} \left(\int_{\underline{\theta}}^{\overline{\theta}} t(\theta) f(\theta) d\theta - c(Q) \right) \text{ s.t.}$$

$$\text{(IC)}\ \ v(\theta, q(\theta)) - t(\theta) \geq v(\theta, q(\theta')) - t(\theta') \text{ for any } \theta,\, \theta'$$

$$\text{(IR)}\ \ v(\theta, q(\theta)) - t(\theta) \geq \overline{u}(\theta) \text{ for any } \theta,$$

where $Q = \int_{\underline{\theta}}^{\overline{\theta}} q(\theta) f(\theta) d\theta$ and, using the additive separability of $c(\cdot)$, $c(Q) = \int_{\underline{\theta}}^{\overline{\theta}} c(q(\theta)) f(\theta) d\theta$. We refer to this model in which the seller's constraints are IC and IR as the *IR model* and define an allocation $\{u(\theta), q(\theta)\}$ to be *implementable* if it satisfies the IC and IR constraints. The IC constraint of a consumer of type θ is satisfied if choosing $q(\theta)$ for the price $t(\theta)$ maximizes the left-hand side of the constraint. Taking first-order conditions, this requires $v_q(\theta, q(\theta)) q'(\theta) = t'(\theta)$ or, equivalently, $u'(\theta) = v_\theta(\theta, q(\theta))$. Since $v_{\theta q}(\theta, q) > 0$, an allocation is incentive compatible if, and only if, it is *locally* incentive compatible in that $u'(\theta) = v_\theta(\theta, q(\theta))$, the schedule $q(\theta)$ is weakly increasing (a.e.), and the utility $u(\theta)$ is absolutely continuous. Since the functions $t(\theta)$ and $q(\theta)$ of an incentive compatible menu are continuous and monotone, we can represent this menu as a *tariff* or price schedule, $T(q)$: the tariff pair $(T(q), q)$ corresponds to the menu pair $(t(\theta), q(\theta))$ evaluated at each θ such that $q = q(\theta)$. We use these *menu* and *tariff* interpretations interchangeably throughout.

Crucial for the characterization of the seller's optimal menu are the seller's first-order condition

$$v_q(\theta, q(\theta)) - c'(Q) = \frac{\gamma(\theta) - F(\theta)}{f(\theta)} v_{\theta q}(\theta, q(\theta)) \tag{1}$$

for each type, and the complementary slackness condition on the IR constraints,

$$\int_{\underline{\theta}}^{\overline{\theta}} [u(\theta) - \overline{u}(\theta)] d\gamma(\theta) = 0. \tag{2}$$

In (1) and (2), $\gamma(\theta) = \int_{\underline{\theta}}^{\theta} d\gamma(x)$ is the cumulative multiplier associated with the IR constraints, which has the properties of a cumulative distribution function, that is, it is nonnegative, weakly increasing, and $\gamma(\overline{\theta}) = 1$. See the Supplementary Appendix for details.[8] Jullien (2000) shows that under the assumptions of *potential*

[8] The integral in the definition of $\gamma(\theta)$ is interpreted as accommodating not just discrete and continuous distributions but also mixed discrete-continuous ones. That is, this formulation allows for the possibility that the IR constraints bind at isolated points. See the Supplementary Appendix for details. It is understood that $q(\theta)$ is evaluated taking the left-limit at jump points.

separation (PS), *homogeneity* (H), and *full participation* (FP), there exists a unique optimal solution to the seller's problem in which all consumers participate, characterized by the first-order conditions (1) and the complementary slackness condition (2) with $q(\theta)$ weakly increasing.

Assumption (PS) effectively ensures that the optimal $q(\theta)$ is weakly increasing. Formally, note that for each type θ, the first-order condition in (1) defines the optimal quantity as a function of the primitives of the economy and the cumulative multiplier $\gamma(\theta)$. It is convenient to define the function $l(\tilde{\gamma}, \theta)$ as the quantity that satisfies (1) at θ when the actual cumulative multiplier $\gamma(\theta)$ is replaced by an arbitrary one, $\tilde{\gamma} \in [0, 1]$. Note that $l(\tilde{\gamma}, \theta)$ weakly decreases with $\tilde{\gamma}$. Assumption (PS) requires $l(\tilde{\gamma}, \theta)$ to weakly increase with θ for all $\tilde{\gamma} \in [0, 1]$, and thus guarantees that the seller has an incentive to discriminate across consumers.[9] Assumption (H) requires that there exist a quantity profile $\{\overline{q}(\theta)\}$ such that the allocation with full participation $\{\overline{u}(\theta), \overline{q}(\theta)\}$ is implementable in that $\overline{u}'(\theta) = v_\theta(\theta, \overline{q}(\theta))$ and $\overline{q}(\theta)$ is weakly increasing. This assumption guarantees that the reservation quantity profile is incentive compatible so that the IC constraints can be satisfied when the IR constraints bind. Assumption (FP) requires that all types participate. Sufficient conditions for (FP) are (H) and $s(\theta, \overline{q}(\theta)) \geq \overline{u}(\theta)$, which guarantees that the seller has an incentive to trade with all consumers.

The solution to the seller's problem is $q(\theta) = l(\gamma(\theta), \theta)$ with associated price $t(\theta) = v(\theta, q(\theta)) - u(\theta)$, where $u(\theta)$ is the utility reached by a consumer of type θ and equals $\overline{u}(\theta)$ for consumer types whose IR constraints bind. As mentioned, a consumer's reservation utility, $\overline{u}(\theta)$, can be equivalently interpreted as the value of purchasing from another seller, producing the good at home, or receiving it as a transfer. By varying the level of the reservation utility, the model can accommodate very different degrees of market power for a seller, ranging from no market power to oligopoly and monopoly power. Specifically, when the reservation utility equals the social surplus under first best for each type, that is, $\bar{u}(\theta) = v(\theta, q_{FB}(\theta)) - c(q_{FB}(\theta))$, the solution to the seller's problem implies $\gamma(\theta) = F(\theta)$ for all consumers so that consumers purchase from the seller first-best quantities, $\{q_{FB}(\theta)\}$, at first-best prices, $\{c(q_{FB}(\theta))\}$. When the reservation utility is lowered from its maximal value under first best, profits correspondingly increase, allowing the model to capture any degree of market power.[10] This feature of the model provides an important dimension of flexibility over the standard model for the measurement exercises in later sections.

To build intuition for how the model works and differs from the standard model, it is useful to consider the case when the degree of convexity of $\overline{u}(\theta)$ is high, the *highly convex* case, and so the IR constraints bind only for isolated types; see Jullien (2000) for details.[11] Note that since $q(\theta)$ is continuous, $\gamma(\theta)$ can have mass points, and thus the IR constraints can bind at isolated points, only at $\underline{\theta}$ or $\overline{\theta}$. In particular, when $\gamma(\theta) = 0$ for $\theta < \bar{\theta}$, the IR constraints bind only at $\overline{\theta}$. In contrast, when $\gamma(\theta) = 1$ for all types, the IR constraints bind only at $\underline{\theta}$ and the model reduces to the standard model, in which the IR constraints simplify

[9]See the proof of Proposition 1 for sufficient conditions on primitives for (PS) to be satisfied and Jullien (2000) for details.

[10]Assumption (H) naturally holds in a model of seller competition with vertical differentiation. Under such interpretation of our model, here we characterize the best-response problem of any such oligopolist. See the Supplementary Appendix.

[11]The monotonicity of $\overline{u}(\theta)$ is an immediate implication of assumption (H), since $\overline{u}'(\theta) = v_\theta(\theta, \overline{q}(\theta))$ and $v_\theta(\cdot, \cdot) > 0$ by assumption. Our analysis would equally apply to the case in which consumers dislike the seller's good and can be ranked by their distaste for it, with θ replaced by $-\theta$, under analogous conditions. We model $\overline{u}(\theta)$ as increasing with θ for consistency with the fact that nearly all households in each village consume the goods we consider. The convexity of $\overline{u}(\theta)$ is to prevent bunching and is implied by the requirement of assumption (H) that $\overline{q}(\theta)$ be weakly increasing.

to $u(\theta) \geq \overline{u}$ with $\overline{u}$ constant. When $\gamma(\theta) = \gamma \in (0, 1)$ for all $\theta \in [\underline{\theta}, \overline{\theta})$, the IR constraints bind at both $\underline{\theta}$ and $\overline{\theta}$. In this case, $F(\theta)$ starts below γ, crosses γ at some type, θ_{HC}, and then lies strictly above it. Types below θ_{HC} consume quantities below first best, whereas types above θ_{HC} consume quantities *above* first best.

3.2 A Model with Heterogeneous Budget Constraints

Suppose now that instead of having heterogeneous outside options, consumers face heterogeneous subsistence constraints. These constraints limit the amount of resources a consumer can spend on a seller's good and formally give rise to a *budget constraint* for the good. We show that under simple conditions, this model and the one of the previous subsection are equivalent in that they imply the same choice of price schedule by a seller and, thus, the same participation and purchase decisions by consumers.

Setup. Suppose that consumers have quasilinear preferences over the seller's good, q, and the numeraire, z, which represents all other goods. A consumer is characterized by a preference attribute, θ, which, as before, affects her valuation of q, and by a productivity attribute, w, which affects her overall budget or income, $Y(w)$.[12] The consumer faces a *subsistence constraint* on the consumption of z of the form $z \geq \underline{z}(\theta, q)$, which can be interpreted as capturing the notion that a certain number of calories are necessary for survival and can be achieved by consuming the seller's good and the numeraire. To see how, define the *calorie constraint* $C^q(\theta, q) + C^z(\theta)z \geq \underline{C}(\theta)$, where $C^q(\theta, q)$ and $C^z(\theta)z$ are, respectively, the calories produced by the consumption of q units of the seller's good and z units of the numeraire for a consumer of type θ, and $\underline{C}(\theta)$ is the subsistence level of calories for such a consumer.[13] Clearly, this calorie constraint can be rewritten as $z \geq [\underline{C}(\theta) - C^q(\theta, q)]/C^z(\theta) \equiv \underline{z}(\theta, q)$.

Let $T(q)$ be the seller's price schedule, where $T(q)$ is the price of quantity q. Conditional on purchasing from the seller, the consumer's problem is

$$\max_{q,z}\{v(\theta, q) + z\} \ \text{ s.t. } \ T(q) + z \leq Y(w) \text{ and } z \geq \underline{z}(\theta, q). \tag{3}$$

Using the fact that the budget constraint holds with equality at an optimum and substituting $z = Y(w) - T(q)$ into the objective function and the constraint $z \geq \underline{z}(\theta, q)$ in (3), the problem in (3) can be restated as

$$\max_{q}\{v(\theta, q) - T(q)\} + Y(w) \ \text{ s.t. } \ T(q) \leq I(\theta, q, w) \equiv Y(w) - \underline{z}(\theta, q), \tag{4}$$

where $I(\theta, q, w)$ is the maximal amount that the consumer can pay to purchase q units of the seller's good and meet her subsistence constraint.[14] Note that the constraint in (4) is a *budget constraint for the seller's good* arising from the consumer's subsistence constraint. We assume that $I(\theta, q, w)$ is absolutely continuous

[12] We implicitly assume that utility is separable across a seller's goods and the seller prices them independently. See Stole (2007) for a discussion and a formalization of the several instances in which this assumption is appropriate.

[13] This formulation of the calorie constraint generalizes $C^q q + C^z z \geq \underline{C}$, used, for instance, by Jensen and Miller (2008), where C^q and C^z are the calories provided by one unit of q and one unit of z, and $\underline{C}$ is the subsistence intake.

[14] Quasilinear preferences in q and z, and so in q and T, are standard in the literature. The more general formulation of preferences as $v(\theta, q, T)$ typically gives rise to a nonconvex constraint set for the seller that renders the characterization of the optimal menu problematic and random tariffs desirable. The latter, however, are unrealistic in the context of our application.

in θ, and twice continuously differentiable and weakly increasing in θ and q.

Suppose that when consumers do not purchase from the seller, they can achieve the exogenous utility level $\overline{u}$, which is constant with θ as in the standard model. Then, the seller's optimal menu solves

$$\text{(BC problem)} \qquad \max_{\{t(\theta),q(\theta)\}} \left(\int_{\underline{\theta}}^{\overline{\theta}} t(\theta) f(\theta) d\theta - c(Q) \right) \text{ s.t.}$$

$$\text{(IC)} \;\; v(\theta, q(\theta)) - t(\theta) \geq v(\theta, q(\theta')) - t(\theta') \text{ for any } \theta, \theta'$$

$$\text{(IR')} \;\; v(\theta, q(\theta)) - t(\theta) \geq \overline{u} \text{ for any } \theta$$

$$\text{(BC)} \;\; t(\theta) \leq I(\theta, q(\theta), w) \text{ for any } \theta.$$

We refer to this model in which the seller's constraints are IC, IR', and BC as the *BC model*, and to an allocation that satisfies these constraints as *implementable*. Although the model admits heterogeneity in both θ and w among consumers, in this section, for expositional simplicity, we consider the case of constant w and suppress the dependence of $I(\theta, q, w)$ and all other variables on w. We examine the implications of this additional dimension of heterogeneity in Appendix A and consider the general case in the empirical analysis, as explained in Section 4.1.

We maintain the same potential separation (PS) and full participation (FP) assumptions as in the IR model. In analogy to assumption (H), we assume there exists an incentive compatible menu $\{\overline{t}(\theta), \overline{q}(\theta)\}$ that induces each consumer to spend her entire budget for the good, that is,

$$\text{(BCH)} \quad \overline{t}(\theta) = I(\theta, \overline{q}(\theta)), \overline{t}'(\theta) = v_q(\theta, \overline{q}(\theta))\overline{q}'(\theta), \text{ and } \overline{q}(\theta) \text{ is weakly increasing.} \tag{5}$$

Importantly, under assumption (BCH), incentive compatibility can be satisfied when the budget constraint $t(\theta) \leq I(\theta, q(\theta))$ binds. As in the IR model, this condition is key to ensuring that there exists an implementable menu that induces all consumers to participate.[15]

Since income affects consumers' purchase behavior in this model, changes in the distribution of income across consumers, due, for instance, to income transfers, typically influence a seller's price schedule. In the IR model, instead, changes in income have *no* impact on the consumption of the seller's good and, thus, on the seller's pricing decisions—unless a consumer's reservation utility, $\bar{u}(\theta)$, is affected in some unspecified way by a change in income. We explore these different implications of the BC and IR models below when we assess the effect of the Progresa transfer on prices.

Equivalence Between Participation and Budget Constraints. The seller's problem with constraints IC, IR', and BC has no known solution. Here we proceed to characterize a seller's optimal menu indirectly by establishing an equivalence between the BC problem and the IR problem. A natural approach, which leads

[15] Sufficient conditions for full participation, and so the IR' constraints, to be satisfied are $v(\theta, \overline{q}(\theta)) - I(\theta, \overline{q}(\theta)) \geq \overline{u}$ and $I(\theta, \overline{q}(\theta)) \geq c(\overline{q}(\theta))$ for each θ. Note first that (BCH) and $I(\theta, \overline{q}(\theta)) \geq c(\overline{q}(\theta))$ ensure that no type is excluded because of a violation of the BC constraint: the seller is better off by offering $\overline{q}(\theta)$ to type θ at price $I(\theta, \overline{q}(\theta))$ for a profit of $I(\theta, \overline{q}(\theta)) - c(\overline{q}(\theta))$ than by excluding this consumer for a profit of zero. Observe next that $v(\theta, \overline{q}(\theta)) - I(\theta, \overline{q}(\theta)) \geq \overline{u}$ guarantees that the IR' constraint is satisfied when the BC constraint binds. Thus, all types participate. Under (FP), the IR' constraints are effectively redundant.

to a simple constructive argument, would be to define the budget for the seller's good of a consumer of type θ as $I(\theta, q(\theta)) = v(\theta, q(\theta)) - \overline{u}_{IR}(\theta)$, where here $\overline{u}_{IR}(\theta)$ denotes the reservation utility of a consumer of type θ in the IR problem. Since, by definition, $t(\theta) = v(\theta, q(\theta)) - u(\theta)$, it is immediate that in this case, the BC constraint is equivalent to the IR constraint of the IR problem. Although this approach is intuitive since it directly relates reservation utilities to budgets, it is unduly restrictive: it requires the schedules of reservation utilities in the IR problem and budgets in the BC problem to agree for each type. For the two problems to admit the same solution, instead, it is sufficient that reservation utilities and budgets, and the derivatives of consumers' utility function and the budget schedule with respect to quantity, agree just for types whose IR constraints bind in the IR problem—as long as consumers have enough income to afford the IR allocation.[16]

Formally, as shown in Appendix A, the BC problem can be conveniently restated as

$$\max_{\{q(\theta)\}} \left(\int_{\underline{\theta}}^{\overline{\theta}} \left\{ v(\theta, q(\theta)) + \left[\frac{F(\theta) - \Phi(\theta)}{f(\theta)} \right] v_\theta(\theta, q(\theta)) + \frac{\phi(\theta) \left[I(\theta, q(\theta)) - v(\theta, q(\theta)) \right]}{f(\theta)} \right\} f(\theta) d\theta - c(Q) \right), \quad (6)$$

with $q(\theta)$ weakly increasing and $u(\theta) \geq \overline{u}$. We term (6) the *simple BC problem*, where $\Phi(\theta) = \int_{\underline{\theta}}^{\theta} \phi(x) dx$ is the cumulative multiplier, defined analogously to $\gamma(\theta)$, on the budget constraint expressed as $I(\theta, q(\theta)) \geq t(\theta) = v(\theta, q(\theta)) - u(\theta)$ and $\phi(\theta)$ is its derivative. The first-order conditions of this problem are

$$v_q(\theta, q(\theta)) - c'(Q) = \left[\frac{\Phi(\theta) - F(\theta)}{f(\theta)} \right] v_{\theta q}(\theta, q(\theta)) + \frac{\phi(\theta) [v_q(\theta, q(\theta)) - I_q(\theta, q(\theta))]}{f(\theta)} \quad (7)$$

for each type, along with the complementary slackness condition

$$\int_{\underline{\theta}}^{\overline{\theta}} \{ I(\theta, q(\theta)) - [v(\theta, q(\theta)) - u(\theta)] \} d\Phi(\theta) = 0. \quad (8)$$

Result 1 in the proof of Proposition 1 states that an implementable allocation is optimal if, and only if, there exists a cumulative multiplier function $\Phi(\theta)$ such that conditions (7) and (8) are satisfied, with $\Phi(\overline{\theta}) = 1$. Next we establish the desired equivalence, denoting by $\{t_{IR}(\theta), q_{IR}(\theta)\}$ the optimal menu in the IR problem.

Proposition 1 (*Equivalence of Problems*)**.** *Assume (PS), (BCH), and (FP). Suppose that the allocation that solves the IR problem is affordable in the BC problem in that $I(\theta, q_{IR}(\theta)) \geq v(\theta, q_{IR}(\theta)) - \overline{u}_{IR}(\theta)$, with equality for types whose IR constraints bind, and $\overline{u}_{IR}(\underline{\theta}) \geq \overline{u}$. If $I_q(\theta, q_{IR}(\theta))$ equals $v_q(\theta, q_{IR}(\theta))$ for types whose IR constraints bind, then the solution to the BC problem coincides with that to the IR problem.*

For intuition, note that in the IR model, a seller in principle can induce a consumer to purchase by offering a high enough quantity for a given price or a low enough price for a given quantity. The IR constraint, however, implicitly places a restriction on the maximal price a seller can charge to a consumer, since the requirement $u_{IR}(\theta) \geq \overline{u}_{IR}(\theta)$ is equivalent to $v(\theta, q_{IR}(\theta)) - \overline{u}_{IR}(\theta) \geq t_{IR}(\theta)$, which effectively limits a

[16] Note that if we replaced (IR') with (IR), it would be easier to establish that the IR and BC problems are equivalent by showing that (IR) implies (BC). We chose a formulation of the BC problem with participation constraints of the form (IR') rather than (IR) to highlight that the equivalence of interest holds more generally, even when participation constraints differ in the two problems.

consumer's expenditure on the seller's good. Hence, in this precise sense, IR and BC constraints are related. Proposition 1 follows by combining this intuition with the construction of a multiplier function on the BC constraints such that the BC constraints bind in the BC problem if, and only if, the IR constraints bind in the IR problem. When this is the case, by comparing (1) and (7), it is easy to see that the first-order conditions of the two problems, and so the optimal quantity schedules, coincide if $v_q(\theta, q_{IR}(\theta))$ equals $I_q(\theta, q_{IR}(\theta))$ for consumers whose IR constraints bind. The first two conditions in the proposition guarantee not just that the solution to the IR problem is feasible for the BC problem but also that utilities, and hence price schedules, in the two problems coincide. This equivalence result extends to any BC problem in which preferences are obtained from an affine transformation of those in the IR problem.

Corollary 1 (*Preferences for Equivalence*)**.** *Consider an IR problem with preferences given by* $v(\theta, q)+z$ *and a BC problem with budget for the seller's good given by* $I(\theta, q)$*. Then, under the conditions of Proposition 1, the solution to the new BC problem with preferences* $\eta_0(\theta) + \eta_1(\theta)[v(\theta, q) + z]$*, with* $\eta_0(\theta)$ *and* $\eta_1(\theta)$ *increasing, and the solution to the IR problem also coincide.*

Proposition 1 and its corollary are important for several reasons. First, these results provide a simple argument for how a model with heterogeneous budget constraints can be represented as a model with heterogeneous reservation utilities, and its solution characterized, even when the preferences in the two problems do not coincide. Second, these results allow us to consider a number of cases of empirical and practical relevance. For instance, we can examine how subsistence constraints affect prices, consumption, and consumers' utilities. We can also evaluate the effect of policies, such as cash transfers, that directly affect consumers' ability to pay and budgets. A natural question, then, is how stringent the assumptions of Proposition 1 are, in particular (BCH), which we address next when $v(\theta, q) = \theta\nu(q)$, a specification of utility common in the literature that we maintain in our empirical analysis.

Proposition 2 (*Subsistence Functions for Equivalence*)**.** *Let utility be* $v(\theta, q) = \theta\nu(q)$ *and the subsistence level be* $\underline{z}(\theta, q) = -\underline{z}_1(\theta) - z_2\nu(q)$ *with* $\underline{z}_1'(\theta) = \psi(\log(\theta - z_2))$*,* $\psi(\cdot)$ *positive and continuous, and* $\underline{\theta} > z_2 > 0$*. Then, (BCH) is satisfied, and* $I_q(\theta, q)$ *equals* $v_q(\theta, q)$ *for types whose BC constraints bind.*

The conditions in Proposition 2 imply that assumption (BCH) holds for a large class of utility functions, $v(\theta, q)$, and subsistence functions, $\underline{z}(\theta, q)$, compatible with Proposition 1; see the proof of Proposition 2 for examples. The assumption that $\underline{z}(\theta, \cdot)$ decreases with q, which is equivalent to assuming that the calorie intake from q, $C(\theta, \cdot)$, increases with q, is natural: the greater the amount of the seller's good consumed, the greater the calorie intake. The requirement that $\psi(\cdot)$ be positive, or, equivalently, that $\underline{z}(\cdot, q)$ decreases with θ, is to ensure that $\overline{q}(\theta)$ is increasing. An intuition for why $\underline{z}(\cdot, q)$ may decrease with θ in practice is that if the same calorie intake can be reached through different combinations of food items, a consumer who values the seller's good more may require *less* of other goods to achieve it. See Lancaster (1966) on the distinction between the caloric and taste attributes of goods, and Jensen and Miller (2008) on the relationship between these attributes and subsistence constraints.

3.3 Nonlinear Pricing: Properties and Implications

By the equivalence just established between models with heterogeneous reservation utilities and heterogeneous budget constraints, from now on we refer to the IR model as the *augmented model* and interpret it as covering both cases. We now examine the implications of the augmented model for prices, consumption, and market participation, for consumer surplus and welfare under nonlinear and linear pricing, and for the impact of policies, such as income transfers, that affect consumers' ability to pay. We maintain assumptions (PS), (H), and (FP), and, for simplicity, that $v(\theta,q)=\theta\nu(q)$ and $c'(Q)=c>0$, as common in the literature. Our results can be extended to the case of nonseparable utility and increasing and convex cost functions.

Prices and Consumption. We start by providing sufficient conditions for quantity discounts to arise and show that quantity discounts are consistent with consumption both below and above first best. Formally, since $q(\theta)$ is increasing, we can define the inverse function $\theta(q)$ and derive the observed price schedule, $T(q) = t(\theta(q))$, as a function of quantity. Using $\theta'(q) = 1/q'(\theta)$, we can then rewrite the local incentive compatibility condition, $\theta\nu'(q(\theta))q'(\theta) = t'(\theta)$, as $\theta\nu'(q(\theta)) = T'(q(\theta))$ so that (1) becomes

$$\frac{T'(q(\theta)) - c}{T'(q(\theta))} = \frac{\gamma(\theta) - F(\theta)}{\theta f(\theta)}. \tag{9}$$

The price schedule $T(q)$ exhibits *quantity discounts* if $T''(q) \leq 0$ or $p'(q) \leq 0$, where $p(q)=T(q)/q$ is the unit price of quantity $q = q(\theta)$; the two characterizations are equivalent when $q(\underline{\theta}) = 0$. Denote by $\varepsilon_r(\theta)$ the elasticity of $F(\theta)/f(\theta)$, which is the reciprocal of the reverse hazard of the distribution of types, with respect to θ, and by $A(\overline{q}(\theta))\equiv-\nu''(\overline{q}(\theta))/\nu'(\overline{q}(\theta))$ the coefficient of absolute risk aversion at $\overline{q}(\theta)$.

Proposition 3 (*Quantity Discounts*)**.** *Assume $\nu'(\cdot) > 0$, $\nu''(\cdot) < 0$, $\frac{\partial}{\partial\theta}\left(\frac{1-F(\theta)}{f(\theta)}\right) \leq 0$, and $\overline{q}'(\theta) > 0$. If $\varepsilon_r(\theta) \geq 1$ and $\overline{u}''(\theta) \geq \nu'(\overline{q}(\theta))/[\theta A(\overline{q}(\theta))]$ for each θ and $\underline{\theta} f(\underline{\theta}) \geq 1$, then $T''(q) \leq 0$ for each q.*

Whereas the first three conditions in the proposition are standard—the third one is a common sufficient condition for assumption (PS)—the remaining ones are novel. The condition $\overline{q}'(\theta)>0$ is a technical assumption to guarantee that $T''(q)$ is bounded when IR constraints bind. The restriction $\varepsilon_r(\theta) \geq 1$ is equivalent to $\frac{\partial}{\partial\theta}\left(F(\theta)/f(\theta)\right) \geq F(\theta)/f(\theta)$ and thus strengthens the usual condition on the distribution of types for assumption (PS) to hold in models with heterogeneous reservation utilities, $\frac{\partial}{\partial\theta}\left(F(\theta)/f(\theta)\right) \geq 0$. It guarantees that the seller has an incentive to discriminate across consumers. The condition that $\overline{u}''(\theta)$ be large enough ensures that consumers whose IR constraints bind are offered quantity discounts. Intuitively, convexity implies that outside consumption opportunities are increasingly more valuable for consumers of higher types. By offering higher quantities at lower marginal prices ($T''(q) \leq 0$), the seller can profitably induce higher types to distinguish themselves from lower types by purchasing more. Then, quantity discounts are optimal for the seller. The condition $\underline{\theta} f(\underline{\theta}) \geq 1$ plays a similar role for types whose IR constraints do not bind.

By comparing the first-order condition in (9) with the one for the first-best allocation, $T'(q(\theta))=c$, it is immediate that the quantity provided to a consumer of type θ is below first best when $\gamma(\theta)>F(\theta)$, whereas it is above first best when $\gamma(\theta) < F(\theta)$. Underprovision arises when the reservation utility is low for high consumer types. In this case, a seller just needs to induce low types to purchase—if they do, high types purchase too—and devise a menu to best extract surplus from the most profitable segment of demand, high consumer types. Since high types anticipate this rent-extraction behavior by sellers, they have an incentive to copy the behavior of lower types and purchase small quantities. But since higher types face a higher marginal benefit from consuming the good, a seller can nonetheless separate them from lower types by decreasing the offered quantities meant for lower types below the lower types' first-best level of consumption. This way, a seller makes the purchase of a small quantity by high-type consumers especially costly for them.

Overprovision, instead, arises when the reservation utility is large for high consumer types. In this case, a seller needs to induce high types to buy in the first place, and can do so while separating them from low types by offering quantities meant for high types that are larger than first best at marginal prices below marginal cost. This way, a seller can not only induce higher types to trade by purchasing these large quantities but also distinguish them from lower types, who naturally prefer smaller quantities. Sorting occurs because lower types would need to purchase much larger quantities than desirable to them, that is, quantities above the first-best level of consumption of higher types, to mimic the behavior of higher types.[17]

Augmented vs. Standard Model. Relative to the standard model, the augmented model gives rise to higher levels of consumption and, correspondingly, lower marginal prices. The reason is that if the cumulative multiplier $\gamma(\theta)$ is different from one, which is its value in the standard model, it must be strictly smaller than one, since it has the properties of a distribution function. But then the first-order condition in (9) immediately implies that a seller has an incentive to provide higher quantities than under the standard model, which the seller can induce consumers to buy only by charging lower marginal prices. Since higher quantities may be offered for a higher tariff, the overall effect on consumers' utility is ambiguous. When the reservation utility in the augmented model is higher than in the standard model, though, consumer surplus is clearly higher in the augmented model.[18] Intuitively, the more attractive outside consumption opportunities are, the more desirable the seller's offered price and quantity must be to induce a consumer to purchase.

Proposition 4 (*Augmented vs. Standard Models*)**.** *Assume full participation under the standard model. The augmented model implies higher consumption and lower marginal prices than the standard model for each consumer type. If $\overline{u}(\underline{\theta}) \geq \overline{u}$, where $\overline{u}$ is the reservation utility in the standard model, then the augmented model also implies higher consumer surplus than the standard model for each consumer type.*

Nonlinear vs. Linear Pricing. A natural question is whether consumers are better off under nonlinear or linear pricing. We argue here that for nonlinear pricing to be preferred, it must typically lead to greater market participation than under linear pricing. Indeed, when all consumers participate under both pricing schemes and nonlinear pricing entails quantity discounts, consumers are better off under linear pricing. Intuitively, linear pricing is preferred when the quantity provided under linear pricing is larger. Perhaps

[17] Formally, incentive constraints are upward-binding for consumer types whose closest marginal type, namely, a type indifferent between purchasing and not, is above them and downward-binding for types whose closest marginal type is below them, leading, respectively, to consumption above and below first best. In the standard model, the only marginal type is the lowest one.

[18] Whenever $|\gamma(\theta) - F(\theta)| \leq 1 - F(\theta)$, the distortion to marginal surplus is smaller in our model than in the standard model.

more surprisingly, consumers prefer linear to nonlinear pricing even when the quantity provided under linear pricing is *smaller*: in this case, a seller who can price discriminate asks for "too high" a price for the greater quantity the seller is willing to provide. In terms of notation, in the linear (monopoly) problem, a seller charges the unit price p_m for any quantity and a consumer of type θ chooses the quantity q to maximize utility $\theta\nu(q)-p_m q$, has demand function $q_m(\theta)=(\nu')^{-1}(p_m/\theta)$, and obtains utility $u_m(\theta)=\theta\nu(q_m(\theta))-p_m q_m(\theta)$.

Proposition 5 (*Nonlinear vs. Linear Pricing*)**.** *Assume full participation under linear pricing. If $p'(q)\leq 0$ at $q=q(\theta)$ and $q_m(\theta)\geq q(\theta)$, or if $T''(q)\leq 0$ at all $q=q(\theta)$, $\gamma(\theta)<1$, and $q(\theta)>q_m(\theta)$, then a consumer of type θ is better off under linear than under nonlinear pricing.*

Naturally, a consumer who is *excluded* from the market under linear pricing but included under nonlinear pricing prefers nonlinear pricing. The next result shows that consumers who have access to generous enough outside consumption possibilities in that $\overline{q}(\theta)>q_{FB}(\theta)$ can indeed be excluded under linear pricing and so are better off under nonlinear pricing.[19] Note that this situation cannot arise when $\gamma(\theta)=1$ for all types, as is the case in the standard model, since $\overline{q}(\theta)\leq q_{FB}(\theta)$ for all types in this case.[20]

Proposition 6 (*Nonlinear vs. Linear Pricing with Exclusion*)**.** *Assume $\nu''(\cdot)<0$. Let $s(\theta,\overline{q}(\theta))\geq\overline{u}(\theta)$ and $\overline{q}(\theta)>q_{FB}(\theta)$ for consumer types in the interval $[\theta',\theta'']$. If there exists $\hat{\theta}\in[\theta',\theta'']$ with $u_m(\hat{\theta})=\overline{u}(\hat{\theta})$, then an interval of consumer types in $[\hat{\theta},\theta'']$ are excluded under linear pricing but included under nonlinear pricing and so are better off under nonlinear than under linear pricing.*

To understand the condition $\overline{q}(\theta)>q_{FB}(\theta)$, note that with $\overline{u}'(\theta)=\nu(\overline{q}(\theta))$, as implied by assumption (H) or (BCH), large values of $\overline{q}(\theta)$ are associated with a rapidly increasing reservation utility profile: outside consumption possibilities are relatively more attractive for consumers of higher types than of lower types. Intuitively, to induce consumers with these generous outside consumption possibilities to participate, a seller needs to offer a low enough marginal price. Since the marginal price is constant and equals the unit price under linear pricing, such a low linear price would lower profits from all consumers for the benefit of including only a few more. Hence, it would not be profitable to include such consumers under linear pricing; see Example 1 in Appendix A for an illustration. Proposition 6 then highlights an efficiency dimension of nonlinear pricing: whenever a seller can include different consumers at different prices, the seller may have an incentive to serve those consumers who would demand unprofitably large quantities under linear pricing.

Income Transfers. The formalization of subsistence and budget constraints allows us to explore how income transfers, such as those implemented by Progresa, affect sellers' incentives to price discriminate, and so prices and consumption, when individuals face subsistence constraints or, more generally, have a limited ability to pay. As we show here, our model implies that income transfers increase consumption but can also lead to an increase in prices, as sellers adjust their menus in response to consumers' greater ability to pay.

[19] The claim requires the existence of a type at risk of exclusion under linear pricing, whose utility equals $\overline{u}(\theta)$ under linear pricing. For instance, this occurs when a consumer's reservation utility equals first-best utility. In this case, under linear pricing, such a consumer is either included, so $u_m(\theta)=\overline{u}(\theta)$, or excluded, in which case again $u_m(\theta)=\overline{u}(\theta)$.

[20] See Corollary 1 in Jullien (2000) for a proof that if $q(\theta)\leq q_{FB}(\theta)$ for all types, then $\overline{q}(\theta)\leq q_{FB}(\theta)$ under assumptions (PS), (H), or (BCH), and (FP).

Intuitively, when consumers are constrained by a budget for the seller's good, changes in their income affect prices by creating an incentive for a seller to extract more surplus. Suppose, for instance, that consumers receive an income transfer that is independent of their characteristics, that is, $\tau(\theta) = \tau > 0$. Such a transfer naturally gives rise to a uniform increase in the price schedule. Why? Since the quantities offered before the transfer are still incentive compatible after the transfer, a seller can offer the same quantities at higher prices without affecting consumers' behavior. The seller then maximizes profits by increasing the price of each quantity by the amount of the transfer. As a result, prices uniformly increase by the amount τ.

Consider now the more interesting case in which the transfer depends on consumers' characteristics, and so affects individual demands differentially. In the villages in our data, transfers depend on household income in that only sufficiently poor households qualify for them and also depend on the number of children in a household. Since poorer households tend to have more children, transfers are larger for poorer households and so effectively progressive in income; see Attanasio et al. (2013). As poorer households also consume less of normal goods, such as those we consider in our application, than richer ones, assuming a transfer such that $\tau'(\theta) \leq 0$ seems then in line with the data, given the one-to-one relationship between types and quantities in our model.

Note that a transfer $\tau(\theta) \geq 0$ amounts to a decrease in the level of consumers' utility when they spend their entire budgets for the seller's good, from $\theta\nu(\overline{q}(\theta)) - I(\theta, \overline{q}(\theta))$ to $\theta\nu(\overline{q}(\theta)) - I(\theta, \overline{q}(\theta)) - \tau(\theta)$, and, importantly, to an increase in its slope in θ when $\tau'(\theta) \leq 0$. Thus, a transfer makes outside consumption opportunities less attractive overall but relatively *more* desirable to higher types than to lower ones. To induce higher types to trade while preserving incentive compatibility, a seller must then offer all consumers higher quantities at lower marginal prices. As a result, the marginal price of the same percentile of quantities, before and after the transfer, *decreases* in response to the transfer. Yet, the price schedule increases for at least some consumers, since the seller can charge higher prices without losing any consumers, as in the case of a uniform transfer. Thus, the positive effect of the transfer on consumption is attenuated by the associated increase in prices.

To formalize these intuitions, we denote by $\{\overline{q}_\tau(\theta)\}$ the incentive compatible quantity profile when consumers spend their entire budgets for the seller's good after the transfer is introduced. For any quantity $q = q(\theta)$ consumed before the transfer, we also define the quantity $\tilde{q}_\tau(q) = q_\tau(\theta)$ consumed after the transfer that corresponds to the same percentile in the distribution of quantities *after* the transfer as the percentile in the distribution of quantities *before* the transfer that q corresponds to. By definition, $\tilde{q}_\tau(q)$ satisfies

$$\int_{\underline{q}_\tau}^{\tilde{q}_\tau(q)} g_\tau(x)dx \left[\int_{\underline{q}_\tau}^{\bar{q}_\tau} g_\tau(x)dx\right]^{-1} = \int_{\underline{q}}^{q} g(x)dx \left[\int_{\underline{q}}^{\bar{q}} g(x)dx\right]^{-1},$$

where $q = q(\theta)$ and $\tilde{q}_\tau(q) = q_\tau(\theta)$ are the quantities consumed before and after the transfer by type θ, and $g(q)$ and $g_\tau(q)$ are, respectively, the probability density functions of quantities before and after the transfer.

Proposition 7 (*Transfers*). *An income transfer $\tau(\theta) > 0$ with $\tau'(\theta) \leq 0$ leads to a higher price schedule with lower marginal prices for each percentile of the distribution of quantities before and after the transfer and to* *a first-order stochastic improvement in the distribution of quantities for: i) all consumer types, if the binding pattern of the BC constraints is the same before and after the transfer; or ii) any interval of consumer types with binding BC constraints either on the entire interval or only at its extremes before and after the transfer, if the binding pattern of the BC constraints differ before and after the transfer.*

Since, by this result, larger quantities are offered at higher prices, the overall effect on the unit price $T(q)/q$ of the good is ambiguous, though. In general, however, the intensity of price discrimination, as measured by the size of quantity discounts, increases with the transfer regardless of its degree of progressivity, except for the consumers of the largest quantities.

Corollary 2. *If $\nu'''(\cdot) \leq 0$ and the transfer $\tau(\theta) \geq 0$ leads to greater consumption for all consumers, then there exist $\theta_{\max} \leq \overline{\theta}$ and a corresponding percentile of the distribution of quantities before and after the transfer, $p_{\max}$, such that the transfer gives rise to greater price discrimination in that $T''_{\tau}(\tilde{q}_{\tau}(q)) \leq T''(q)$ for each percentile of the distribution of quantities before and after the transfer up to $p_{\max}$.*

Although the exact shape of unit prices depends on the distribution of types and quantities and on how this latter distribution is affected by the transfer, asymmetric changes in unit prices for low and high quantities occur naturally. An example is when $\nu(q)$ is a HARA function with decreasing absolute risk aversion and types are uniformly distributed. In this case, unit prices *increase* at low quantities but *decrease* at high quantities in response to an increase in income—provided utility is sufficiently convex in θ when the BC constraints bind. See the Supplementary Appendix. Hence, the schedule of unit prices becomes steeper. We explore the extent to which these implications of our model are borne out in the data in Section 5.4.

Observe, finally, that an opposite logic applies to the case of in-kind transfers, when they lead to an increase in the consumption floor $\underline{z}(\theta, q)$. Specifically, by reversing the argument that establishes Proposition 7, it is possible to show that an in-kind transfer can lead to an increase in the consumption floor on other goods, which reduces consumers' budgets for the good priced nonlinearly and so gives rise to a decrease in its price. In this case, in-kind transfers would have an *opposite* effect to that of cash transfers and generate a decrease rather than an increase in prices. Consistently with this implication of our model but using Mexican data different from ours, Cunha et al. (2017) indeed show that the price effects of in-kind transfers are nonnegligible and *negative* in more remote and less competitive areas but are otherwise small. In these cases, in-kind transfers may be preferable to cash transfers.

4 Identification and Estimation

In this section, we discuss the identification and estimation of the model's primitives. The basic idea, which builds on intuitions from Perrigne and Vuong (2010), is that the pricing behavior of a seller depends on the property of the distribution of consumer types in a village. Since this distribution can be mapped into the distribution of quantities, it can be recovered from the joint distribution of observed prices and quantities. Although one can show that the parameters of the model can be identified and estimated semiparametrically, here we derive estimators that use flexible parametric functions, partially to accommodate the sparsity of the data in some of the villages. Omitted details and proofs are collected in Appendix B.

4.1 Identification

In a village market, the model's primitives are the consumers' utility function, $v(\theta, q)$, the cumulative distribution function of consumers' types or marginal willingness to pay, $F(\theta)$, its support $[\underline{\theta}, \overline{\theta}]$, and the associated probability density function, $f(\theta)$, the seller's marginal cost at the total quantity provided of the good considered, $c'(Q)$, and the determinants of participation in the market: the reservation schedule $\overline{u}(\theta)$ in the IR model and the budget schedule $I(\theta, q, w)$ in the BC model. We consider the general version of the BC model with heterogeneity in θ and w, both of which are assumed to be noncontractible. We allow for dependence between θ and w so that, without loss of generality, we can interpret the budget schedule as a function of θ only, with $\Upsilon(\theta) \equiv I(\theta, q(\theta), \omega(\theta))$; see the discussion of the two-dimensional case in Appendix A.[21] Under standard assumptions, we show that these primitives are identified in each village from data on consumers' expenditures and quantity purchases, which provide information, respectively, about $T(q)$ and q. Note that $\overline{u}(\theta)$ and $\Upsilon(\theta)$ are only identified for households with binding IR and BC constraints, respectively.[22] Key properties of the model depend on the cumulative multiplier $\gamma(\theta)$, which we refer to simply as the *multiplier*.

In establishing identification, we maintain that the sufficient condition $s(\theta, \overline{q}(\theta)) \geq \overline{u}(\theta)$ for full participation holds: it states that a seller obtains nonnegative profits from each consumer's type at the reservation quantity, $\overline{q}(\theta)$. This approach is justified by the fact that nearly all households in each village consume the three goods we consider, namely, rice, kidney beans, and sugar, as discussed. We also adopt the normalization $\underline{\theta} = 1$, since a scaling assumption on types is necessary for identification. Let $G(q)$ denote the cumulative distribution function of quantities in a village and $g(q)$ the corresponding density function. Since $G(q)$, $g(q)$, the price schedule $T(q)$, and its derivatives are identifiable from information on quantities and prices, we treat them as known in our identification arguments.

Our arguments rely on the condition for local incentive compatibility of an optimal menu, $T'(q) = v_q(\theta, q)$, and a seller's first-order condition for the choice of quantity. This latter condition provides the only condition for the identification of a seller's cost structure. By relying exclusively on information on quantities and prices, we can only identify a seller's marginal cost at the total quantity of a good provided in a village, $Q = \int_{\underline{\theta}}^{\overline{\theta}} q(\theta) f(\theta) d\theta$. Yet, based on this information alone, we can identify all primitives under the assumption that $v(\theta, q) = \theta \nu(q)$, which we maintain from now on, up to the coefficient of absolute risk aversion.[23] This

[21] Note that consumers' marginal willingness to pay and absolute ability to pay, captured here by θ and w, are highly correlated in the data, as apparent from the observed strong relationship between consumption and income and the fact that the commodities we consider are normal goods. Recall also that the budget $I(\theta, q(\theta), \omega(\theta))$ matters only for consumers whose budget constraints bind, in which case it just equals $T(q)$.

[22] Any economy with reservation utility $\overline{u}(\theta)$ or budget schedule $\Upsilon(\theta)$ binding on the set $\Theta' \subseteq \Theta$ is observationally equivalent to an economy with the same primitives but reservation utility $\widetilde{u}(\theta)$ or budget schedule $\widetilde{\Upsilon}(\theta)$ that agree, respectively, with $\overline{u}(\theta)$ and $\Upsilon(\theta)$ on Θ' and are appropriately adjusted for the remaining types.

[23] Restrictions on the utility function are common in the auction and nonlinear pricing literature. Note that in auction models with risk-averse bidders, even restricting the utility function to belong to well-known families of risk aversion is typically not sufficient for identification; see Campo et al. (2011). Here we presume that the absolute risk aversion coefficient is known, but we do not otherwise restrict consumers' utility function or type distribution. When $v(\theta, q)$ is not multiplicatively separable in θ, then $\gamma(\theta)$ is set-identified, but $v_q(\theta, q)$ is still point-identified; see the Supplementary Appendix for this argument and a discussion of related results in the nonlinear pricing and hedonic pricing literature.

multiplicative specification of utility is ubiquitous in the theoretical and empirical literature on auctions and nonlinear pricing for its tractability (see Guerre et al. (2000) and Perrigne and Vuong (2010)). Thus, we consider it a natural benchmark.

Marginal Cost and Multipliers on Constraints. The relationship between θ and q implied by incentive compatibility is central to the identification of the model. To see why, let $\underline{q} \equiv q(\underline{\theta})$ and $\overline{q} \equiv q(\overline{\theta})$ denote, respectively, the smallest and largest observed quantities in a village. Since $q'(\theta) \geq 0$, it follows that $G(q) = F(\theta)$: the cumulative distribution function of types is identified by that of quantities and $g(q) = f(\theta)/q'(\theta)$. Given this mapping between the distribution of types and quantities, a seller's first-order condition can be used to identify the marginal cost $c'(Q)$, the multiplier $\gamma(\theta(q))$ on participation (or budget) constraints, and so the set of consumers whose participation (or budget) constraints bind. Formally, rewrite (9) as

$$\frac{g(q)}{\varphi(q)}\left[\frac{c'(Q)}{T'(q)} - 1\right] = G(q) - \gamma(\theta(q)), \tag{10}$$

with $\varphi(q) \equiv \partial \log(\theta(q))/\partial q = \theta'(q)/\theta(q)$. We show next that both $c'(Q)$ and $\gamma(\theta(q))$ are identified up to the coefficient of absolute risk aversion. As a preliminary step, we show that $c'(Q)$ is identified up to the ratio $\varphi(\overline{q})/\varphi(\underline{q})$. To this purpose, note that, by taking derivatives of both sides of (10) and then integrating the resulting expressions from $\underline{q}$ and $\overline{q}$, it is easy to show that

$$c'(Q) = \left[g(\overline{q})\frac{\varphi(\underline{q})}{\varphi(\overline{q})} - g(\underline{q})\right] / \left[\frac{g(\overline{q})}{T'(\overline{q})}\frac{\varphi(\underline{q})}{\varphi(\overline{q})} - \frac{g(\underline{q})}{T'(\underline{q})}\right];$$

see the proof of Proposition 8 in Appendix A for details. Hence, since $g(q)$ and $T'(q)$ are identified, $c'(Q)$ is identified up to $\varphi(\overline{q})/\varphi(\underline{q})$. To complete the argument, recall that $A(q) = -\nu''(q)/\nu'(q)$ is the coefficient of absolute risk aversion. Differentiating the local incentive compatibility condition $T'(q) = \theta(q)\nu'(q)$ gives $\theta'(q)/\theta(q) = T''(q)/T'(q) + A(q)$. Therefore, condition (10) also implies

$$\gamma(\theta(q)) = G(q) + g(q)\left[1 - \frac{c'(Q)}{T'(q)}\right]\left[\frac{T''(q)}{T'(q)} + A(q)\right]^{-1}. \tag{11}$$

Then, the multiplier $\gamma(\theta(q))$ is identified up to $c'(Q)$ and $A(q)$, given that $G(q)$, $g(q)$, $T'(q)$, and $T''(q)$ are identified. However, since $\varphi(\overline{q})/\varphi(\underline{q})$ only depends on $\theta'(q)/\theta(q)$, which is known up to $A(q)$, it follows that $c'(Q)$ is identified if $A(q)$ is known. Hence, $\gamma(\theta(q))$ is identified just up to $A(q)$. Equation (11) makes it clear that the identification of $c'(Q)$ and the multiplier requires some knowledge of the shape of the utility function. In estimation, we sidestep this issue by specifying $\gamma(\theta(q))$ as a parametric yet flexible function of q, which reduces the identification problem to the problem of identifying $c'(Q)$ and the parameters of this function.

Proposition 8. *In a village, the marginal cost of the total quantity provided, $c'(Q)$, and the schedule of multipliers, $\gamma(\theta(q))$, are identified up to the coefficient of absolute risk aversion. In particular, up to this coefficient, $\gamma(\theta(q))$ is identified from the cumulative distribution function of quantities, $G(q)$, the associated* *probability density function, $g(q)$, and the marginal price schedules $T'(q)$ and $T''(q)$.*

There are special cases of interest in which identification is simpler. For instance, when the reservation utility is highly convex, $\gamma(\theta(q))=\gamma$ is constant on $[\underline{\theta},\overline{\theta})$ so that γ equals $G(q)$ only once between $\underline{q}$ and $\overline{q}$. In this case, $c'(Q)$ and this constant γ are identified just up to one scale normalization, namely, $\varphi(\overline{q})/\varphi(\underline{q})$: once $c'(Q)$ is identified, this constant γ is identified by the value of $G(q)$ at the quantity at which $T'(q)$ equals $c'(Q)$ by (10). When the resulting γ equals one, our model reduces to the standard model.

Distribution of Consumer Types. We now show that the type support, $\theta(q)$, and the probability density function of types, $f(\theta)$, are identified. Note that condition (10) can be rewritten as $\theta'(q)/\theta(q)=g(q)[T'(q)-c'(Q)]/\{T'(q)[\gamma(\theta(q))-G(q)]\}$, which can be used to express $\theta(q)$ as

$$\log(\theta(q)) = \log(\theta(\underline{q})) + \int_{\underline{q}}^{q} \frac{\partial \log(\theta(x))}{\partial x} dx = \log(\theta(\underline{q})) + \int_{\underline{q}}^{q} \frac{g(x)[T'(x) - c'(Q)]}{T'(x)[\gamma(\theta(x)) - G(x)]} dx, \qquad (12)$$

where the first equality in (12) follows by integrating $\partial \log(\theta(q))/\partial q$ with respect to quantity.[24] Once $c'(Q)$ and $\gamma(\theta(q))$ are identified, expression (12) implies that $\theta(q)$ is identified too up to $\theta(\underline{q})$, since it is a known function of objects that are either identified or known, that is, $\underline{q}$, q, $g(q)$, $T'(q)$, and $G(q)$. Then, $f(\theta)$ is identified from $g(q)$ and the derivative $\theta'(q)$, since $f(\theta)=g(q)/\theta'(q)$ by $F(\theta)=G(q)$.

Proposition 9. *In a village, the support of consumers' marginal willingness to pay, $\theta(q)$, is identified from the cumulative distribution and the probability density functions of quantities, $G(q)$ and $g(q)$, the marginal cost $c'(Q)$, the marginal price schedule, $T'(q)$, and the schedule of multipliers, $\gamma(\theta(q))$, up to a scale normalization. The probability density function of consumers' marginal willingness to pay, $f(\theta)$, is identified from the probability density function of quantities, $g(q)$, and the first derivative of $\theta(q)$.*

Utility Function and Schedule of Reservation Utility. Here we show how $\nu'(q)$, $\nu(q)$, and $\overline{u}(\theta)$ are identified. Note first that with knowledge of the coefficient of absolute risk aversion, base marginal utility $\nu'(q)$ is identified up to a scale normalization. Here we show that this scale normalization is unnecessary to identify $\nu'(q)$ if the marginal price schedule, $T'(q)$, and the type support, $\theta(q)$, are identified. To this end, observe that once $\theta(q)$ is identified, the incentive compatibility condition $\nu'(q) = T'(q)/\theta(q)$ implies that $\nu'(q)$ is identified from $T'(q)$. Next, given $\nu'(q)$, we can recover $\nu(q)$ up to one point, say, $q'=q(\theta')$, since $\nu(q)=\nu(q')-\int_{q}^{q'} \nu'(x)dx$ for $q \leq q'$ and $\nu(q)=\nu(q')+\int_{q'}^{q} \nu'(x)dx$ for $q \geq q'$. With $\theta(q)$ and $\nu(q)$ identified, $\overline{u}(\theta)$ is identified for all consumers whose participation (or budget) constraints bind, since their utility is $\overline{u}(\theta)=\theta(q)\nu(q)-T(q)$.

Proposition 10. *In a village, the base marginal utility function, $\nu'(q)$, is identified from the marginal price schedule, $T'(q)$, and the support of consumers' marginal willingness to pay, $\theta(q)$. Hence, $\nu(q)$ is identified*

[24]The integrand in (12) is positive since $g(q)>0$, $T'(q)>0$, and $T'(q)\geq c'(Q)$ if, and only if, $\gamma(\theta(q))\geq G(q)$ by (10). It is well-defined when $\gamma(\theta(q))=G(q)$ and $T'(q)=c'(Q)$ if the slope of $\gamma(\theta(q))$ differs from $g(q)$ at such quantities, which is the case under our assumptions, as apparent from the seller's first-order condition expressed as $\gamma(\theta(q))=G(q)+f(\theta(q))s_q(\theta(q),q)/v_{\theta q}(\theta(q),q)$. To see where this slope condition on $\gamma(\theta(q))$ comes from, denote by q^s any quantity such that $\gamma(\theta(q))=G(q)$ and $T'(q)=c'(Q)$, and note $\lim_{q\to q^s} g(q)[T'(q) - c'(Q)]/\{T'(q)[\gamma(\theta(q)) - G(q)]\}=g(q^s)T''(q^s)/\{T'(q^s)[\gamma'(\theta(q^s))\theta'(q^s) - g(q^s)]\}$.

up to a scale normalization. The reservation utility (or budget) function is identified for all consumers whose participation (or budget) constraints bind.

4.2 Estimation

For each of the three goods considered, we estimate the model separately in each village in two steps. In the first step, we parameterize the functions $T(q)$, $G(q)$, and $\gamma(q)$—this latter as a flexible cumulative distribution function—and estimate them by maximum likelihood together with the model's primitives $c'(Q)$, $\theta(q)$, and $\nu'(q)$. Specifically, we use flexible parameterizations for $T(q)$ and $G(q)$, which, together with the seller's first-order condition, provide the estimating equations for the parameters of $T(q)$, $G(q)$, and $\gamma(\theta(q))$, and for $c'(Q)$. We estimate $\theta(q)$ and $\nu'(q)$ as known transformations of $T'(q)$, $G(q)$, $\gamma(\theta(q))$, and $c'(Q)$ based on equation (12) for $\theta(q)$ and the local incentive compatibility condition $\nu'(q) = T'(q)/\theta(q)$ for $\nu'(q)$. In the second step, we estimate $f(\theta)$ in each village from the estimated $\theta(q)$ via a kernel density estimator. Since the rate of convergence of the parametric estimator of $\theta(q)$ is faster than that of the nonparametric estimator of $f(\theta)$, the estimation of $\theta(q)$ in the first step does not affect this second step.[25] This approach is appealing not only for its simplicity but also for its consistency with the nature of our data, which would make a full nonparametric approach problematic.

Price Schedule and Distribution of Quantities. For each of the three goods we consider, our data contain information on the quantities purchased and the unit prices paid by each household in each village. Denote by N_{vj} the number of consumers of good j in village v and by q_{vji} the quantity of good j purchased by household i in the village. We estimate the price schedule, $\{T_{vj}(q)\}$, of good j in village v as

$$\log[T_{vj}(q_{vji})] = t_{vj0} + t_{vj1}\log(q_{vji}) + \varepsilon^{p}_{vji}, \tag{13}$$

where $T_{vj}(q_{vji}) \equiv E[p(q_{vji})|q_{vji}]q_{vji}$, given the observed unit price $p(q_{vji})$ of quantity q_{vji}, and ε^{p}_{vji} is measurement error in expenditure. We use the mean unit value of a quantity in each village as the unit price of that quantity for consistency with our model: multiple unit values may correspond to a given quantity in a village but our model implies that the price schedule is a *function* of quantity rather than a correspondence. We also do so in order to minimize measurement error in unit values, due, for instance, to recall and recording error—the assumption implicit in (13) is that unit values rather than quantities are contaminated by error. We treat quantity as exogenous, since information on the quantity purchased and the price paid by each household provides direct information on the price schedule of the seller and $T(q)$ is a deterministic function of q in our model. Let $T'_{vj}(q) \equiv \partial E[T_{vj}(q)|q]/\partial q$.

We parameterize the cumulative distribution function of the quantities of good j purchased in village v

[25] We estimated the first-step system also by GMM, but convergence was problematic in many villages due to the sparseness of the data. If $f(\theta)$ is interpreted as the probability mass function associated with the empirical $G(q)$, the second step is unnecessary.

as a logistic function with a quadratic index allowing for error in recorded purchase frequency, ϵ_{vji},

$$G_{vj}(q_{vji}) = \frac{\exp\{g_{vj0} + g_{vj1}q_{vji} + g_{vj2}q_{vji}^2 + \epsilon_{vji}\}}{1 + \exp\{g_{vj0} + g_{vj1}q_{vji} + g_{vj2}q_{vji}^2 + \epsilon_{vji}\}}, \tag{14}$$

where $\{G_{vj}(q_{vji})\}$ is the empirical cumulative distribution function of good j in village v. Note that ϵ_{vji} captures not only recall error but also the error resulting from the timing of the Progresa interview relative to the timing of a household's regular purchases, which may lead to understating or overstating the fraction of households purchasing a particular quantity.

Marginal Cost and Multiplier Function. We rewrite the seller's first-order condition in (10) to relate the cumulative distribution function of quantities, $G(q)$, to marginal cost, $c'(Q)$, and the multiplier, $\gamma(\theta(q))$, as

$$G(q) = \left[\frac{1}{T'(q)} - \frac{1}{c'(Q)}\right] x_3(q) + \gamma(\theta(q)) = \left[\frac{1}{t_1 p(q)} - \frac{1}{c'(Q)}\right] x_3(q) + \gamma(\theta(q)), \tag{15}$$

where $x_3(q) \equiv c'(Q)g(q)\theta(q)q'(\theta)$ and the second equality in (15) follows from (13), which implies $p(q) = T'(q)/t_1$ by dropping subscripts. Denote the marginal cost of the total quantity of good j in village v by $c'_{vj}(Q_{vj})$. We specify the auxiliary function $x_3(q)$ as a (positive) fractional polynomial with noninteger powers, $x_{vj3}(q_{vji}) = \chi_{vj0} + \chi_{vj1}q_{vji}^{\chi_{vj2}}$, to encompass a wide range of shapes. Given the limited granularity of our data, estimating $x_3(q)$ nonparametrically would be infeasible. Since the multiplier function $\gamma(\theta(q))$ has the properties of a cumulative distribution function, we specify it as a flexible logistic function of quantity,

$$\gamma_{vj}(q_{vji}) = \frac{\exp\{\gamma_{vj0} + \gamma_{vj1}q_{vji} + \gamma_{vj2}q_{vji}^2\}}{1 + \exp\{\gamma_{vj0} + \gamma_{vj1}q_{vji} + \gamma_{vj2}q_{vji}^2\}}. \tag{16}$$

We consider both a *linear* specification with γ_{vj2} set to zero and a *quadratic* one, in which we estimate γ_{vj2}. In this latter case, expression (15) becomes

$$\begin{aligned} G_{vj}(q_{vji}) &= \left[\frac{1}{p_{vj}(q_{vji})} - \frac{1}{\underline{c}'_{vj}(Q_{vj})}\right] (\underline{\chi}_{vj0} + \underline{\chi}_{vj1} q_{vji}^{\chi_{vj2}}) + \frac{\exp\{\gamma_{vj0} + \gamma_{vj1}q_{vji} + \gamma_{vj2}q_{vji}^2\}}{1 + \exp\{\gamma_{vj0} + \gamma_{vj1}q_{vji} + \gamma_{vj2}q_{vji}^2\}} + \varepsilon^s_{vji} \\ &= -\frac{\underline{\chi}_{vj0}}{\underline{c}'_{vj}(Q_{vj})} + \frac{\underline{\chi}_{vj0}}{p_{vj}(q_{vji})} - \frac{\underline{\chi}_{vj1} q_{vji}^{\chi_{vj2}}}{\underline{c}'_{vj}(Q_{vj})} + \frac{\underline{\chi}_{vj1} q_{vji}^{\chi_{vj2}}}{p_{vj}(q_{vji})} + \frac{\exp\{\gamma_{vj0} + \gamma_{vj1}q_{vji} + \gamma_{vj2}q_{vji}^2\}}{1 + \exp\{\gamma_{vj0} + \gamma_{vj1}q_{vji} + \gamma_{vj2}q_{vji}^2\}} + \varepsilon^s_{vji}, \end{aligned} \tag{17}$$

with $\underline{c}'_{vj}(Q_{vj}) \equiv c'_{vj}(Q_{vj})/t_{vj1}$, $\underline{\chi}_{vj0} \equiv \chi_{vj0}/t_{vj1}$, and $\underline{\chi}_{vj1} \equiv \chi_{vj1}/t_{vj1}$, and ε^s_{vji} accounts for measurement (or specification) error. Equation (17) provides the estimating equation for marginal cost and the multiplier function. Note that the second line of (17) is the sum of a linear-in-parameters function (the first two terms), a fractional polynomial (the third and fourth terms), and a function of $e^{-q_{vji}}$ and $e^{-q^2_{vji}}$, whose parameters are identified.[26] We test whether the estimated parameters of $\{\gamma_{vj}(q_{vji})\}$ in each village are individually and

[26] From the fact that the multiplier is known at the smallest and largest quantities (from the density function of quantities) and must equal $G(q)$ at least at one interior quantity, at which marginal cost is identified by $T'(q)$, we obtain four moment conditions to identify the three parameters of $\gamma_{vj}(\cdot)$ as well as marginal cost in each village. Then, the identification of the parameters of $x_{vj3}(\cdot)$ based on (17) follows by standard arguments. Note that the empirical $G(\cdot)$ is a step function, so $\lim_{q \downarrow \underline{q}} G(q) = G(q_{vj1})$

jointly significant to infer the shape of the multiplier, and thus the binding pattern of the relevant constraints. Specifically, if γ_{vj1} and γ_{vj2} are not significantly different from zero, we infer that the multiplier is constant, and so, depending on its value, the relevant constraints bind only at $\underline{\theta}$ or $\overline{\theta}$ or at both types. Otherwise, if γ_{vj1} and γ_{vj2} are significantly different from zero, we determine the intervals of types whose constraints bind by testing whether $\gamma_{vj}(q_{vji+1})$ is significantly different from $\gamma_{vj}(q_{vji})$ at any q_{vji} and q_{vji+1}. If not, we infer that the derivative of the multiplier is zero over the corresponding interval of types, and so the constraints do not bind.

Support of Consumer Types and Utility Function. Normalizing $\underline{\theta}$ to 1, we specify consumers' log marginal willingness to pay in village v for good j by (12) as

$$\log(\theta_{vj}(q)) = \frac{1}{N_{vj}} \sum_{i=1}^{N_{vj}} \left(\frac{[T'_{vj}(q_{vji}) - c'_{vj}(Q_{vj})]\mathbf{1}\{q_{vji} \leq q\}}{T'_{vj}(q_{vji})[\gamma_{vj}(q_{vji}) - G_{vj}(q_{vji})]} \right),$$

where N_{vj} is the number of consumers of good j in village v, Q_{vj} is the total quantity purchased in village v, $T'_{vj}(q_{vji})$ is computed from (13), and $G_{vj}(q_{vji})$ and $\gamma_{vj}(q_{vji})$ are specified in (14) and (16). Using local incentive compatibility, $\nu'(q) = T'(q)/\theta(q)$, and the forms of $\theta(q)$ and $T'(q)$, we estimate $\nu'(q)$ as

$$\log(\nu'_{vj}(q)) = \log(T'_{vj}(q)) - \frac{1}{N_{vj}} \sum_{i=1}^{N_{vj}} \left(\frac{[T'_{vj}(q_{vji}) - c'_{vj}(Q_{vj})]\mathbf{1}\{q_{vji} \leq q\}}{T'_{vj}(q_{vji})[\gamma_{vj}(q_{vji}) - G_{vj}(q_{vji})]} \right).$$

Probability Density Function of Types. Given the estimated $\theta_{vji} = \theta_{vj}(q_{vji})$, we estimate the density of households' marginal willingness to pay for good j in village v as $f_{vj}(\theta) = (N_{vj}h^{\theta}_{vj})^{-1} \sum_{i=1}^{N_{vj}} K^{\theta}_{vj}((\theta - \theta_{vji})/h^{\theta}_{vj})$, with Epanechnikov kernel function $K^{\theta}_{vj}(\cdot)$ and bandwidth h^{θ}_{vj}.

5 Empirical Results

In this section, we present the main results obtained by estimating the model on data from the Mexican villages that were part of the survey collected for the evaluation of Progresa, as explained in Section 2. We start by discussing the sample selection criteria used in estimation. We then present estimates of the primitives of the model and discuss the fit of the model to the data. Since we consider many villages, we represent graphically the point estimates of the objects of interest and report their associated t-statistics in Appendix B. Next, we use the model to analyze the distortions implied by the price discrimination we observe and evaluate the welfare implications of alternative pricing schemes. Finally, we derive a reduced form of the first-order conditions for the optimality of sellers' and consumers' behavior that relates unit prices to quantities and the hazard rate of the distribution of quantities in each village. We use this reduced form to explain the effect of the Progresa transfers on prices, in particular on the intensity of price discrimination, which we estimate by exploiting the experimental variation in the data induced by the introduction of the Progresa transfer in a randomly selected subset of villages. Omitted details are collected in Appendix B.

and $\lim_{q \uparrow \overline{q}} = G(q_{vjN_{vj}-1})$, whereas $\gamma_{vj}(\cdot)$ ranges from zero to one. Thus, we do not impose that $G(\cdot)$ equals the multiplier at any quantity.

5.1 Sample Selection

We estimate the model for three commodities—rice, kidney beans, and sugar—which we chose for three reasons, as explained in Section 2. First, they are commonly consumed, so that the full participation assumption is likely to be valid, and we observe a large number of transactions for them. Second, they are normal goods, whose consumption increases with income, so that assuming that households' marginal willingness to pay and absolute ability to pay are related, as we do in our BC model, is plausible. Third, they are goods of homogeneous quality, so that the variation in prices across quantities we document is likely to reflect just quantity discounts. We use five waves of the Progresa survey, adjusting the observed prices over time for inflation. To minimize the impact of measurement error, we exclude extreme observations, both in terms of quantity and, importantly, in terms of prices—we drop the top 5% of observations on (unit) prices and quantities. We also exclude observations reported in incorrect units—say, liters rather than kilos. We focus on villages with at least 75 observations on each of the goods of interest and with at least 50% of unit prices declining with quantity. These restrictions imply the loss of very few observations: the original sample of 191 municipalities is reduced to 170 for rice, 176 for kidney beans, and 183 for sugar.

While in this section a village is defined as a municipality, in Appendix B, we also present the estimates of our model on villages defined as Mexican localities. Applying to this sample of localities the same selection rules applied to the sample of municipalities restricts the original sample of 506 localities to 371 for rice, 396 for kidney beans, and 442 for sugar. The estimates of the model are very similar across villages defined as municipalities and localities.[27]

5.2 Estimates

The core elements of our model in each village are the unit price schedule $p(q)$, the cumulative distribution function of quantities $G(q)$, which is mapped in the cumulative distribution function of types $F(\theta)$, the multiplier $\gamma(\theta(q))$, the type support, $\theta(q)$, and consumers' utility function, $\theta(q)\nu(q)$. In this subsection, we first present the estimates of these objects in each municipality and then illustrate the fit of the model to the data. See Appendix B for the omitted t-statistics of these estimates and the Supplementary Appendix for the estimates of the probability density function of consumer types.

Estimation Sample: Prices and Quantities. As discussed in Section 4.2, we estimate both a linear and a quadratic version of the logistic index of the multiplier function in equation (16). For the linear specification of the multiplier, we successfully estimate the model primitives for 134, 144, and 143 of the 170, 176, and 183 municipalities considered in the estimation sample for rice, kidney beans, and sugar, respectively. For the quadratic specification, we obtained estimates for 142, 153, and 139 of these municipalities.[28] For this sample, in the top panels of Figure 1 we report the schedule of mean unit prices per quantity in each village,

[27] Restricting attention to villages with at least 50% of unit prices declining with quantity accounts for a very small loss of villages, namely, 3, 8, and 3 villages defined as municipalities, and 2, 31, and 13 villages defined as localities, respectively, for rice, kidney beans, and sugar. This loss is primarily due to the irregularity of the price schedule in these excluded villages.

[28] For the remaining villages defined as either municipalities or localities, we could not obtain estimates of the primitives due to numerical issues—the estimation routine failed to achieve convergence—largely stemming from the sparsity of the data.

computed as explained above, together with a fractional polynomial interpolating line, the solid line. In the bottom panels of Figure 1, we plot the cumulative distribution function of quantities in each village. All goods are measured in kilos. As apparent, in most villages the unit price of each good declines with quantity, so unit prices are highest for the households who purchase the smallest quantities and decrease more rapidly over the range of small quantities that most households purchase, as evident from comparing the top panels and bottom panels. Thus, households are affected by the nonlinearity of prices, and most of them face significant quantity discounts. For instance, the mean unit price of the smallest quantity of rice, 0.1 kilos, in general is more than 5 pesos, whereas that of the largest quantity, 2 kilos, can be as low as 1 peso.

Figure 1: Unit Prices and Cumulative Distribution Function of Quantities

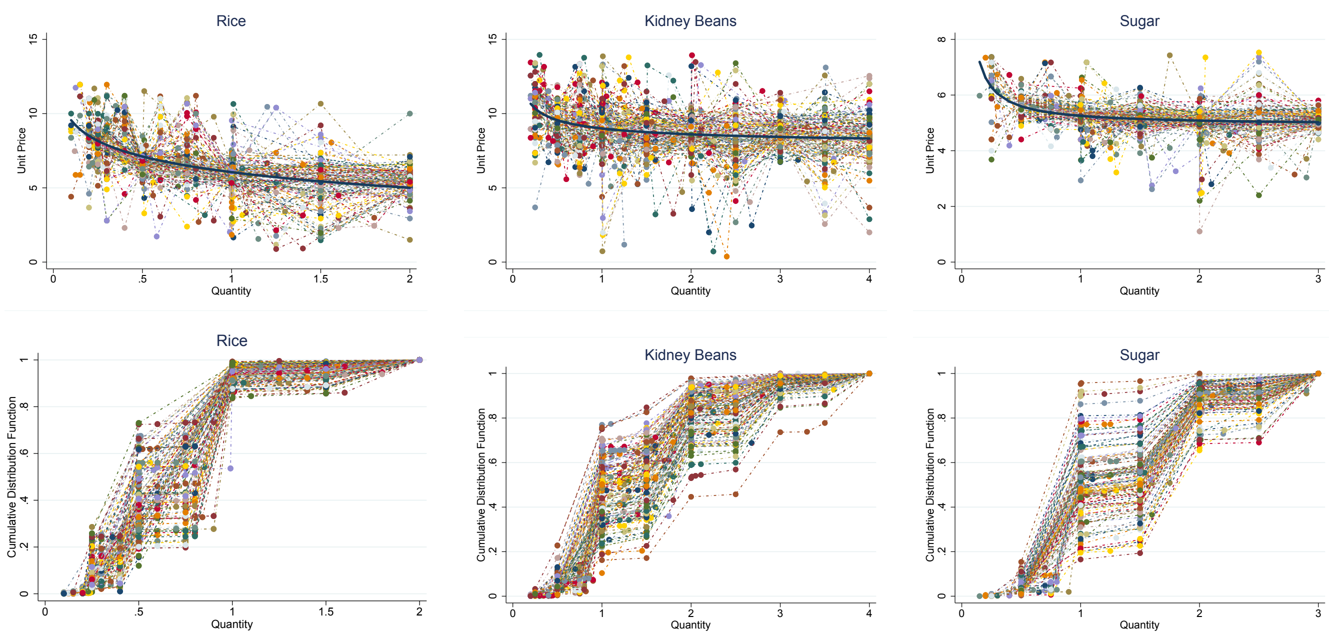
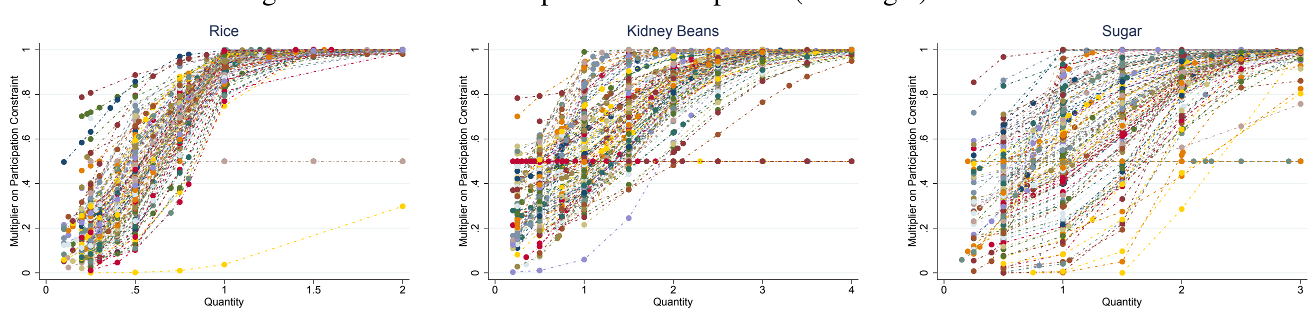


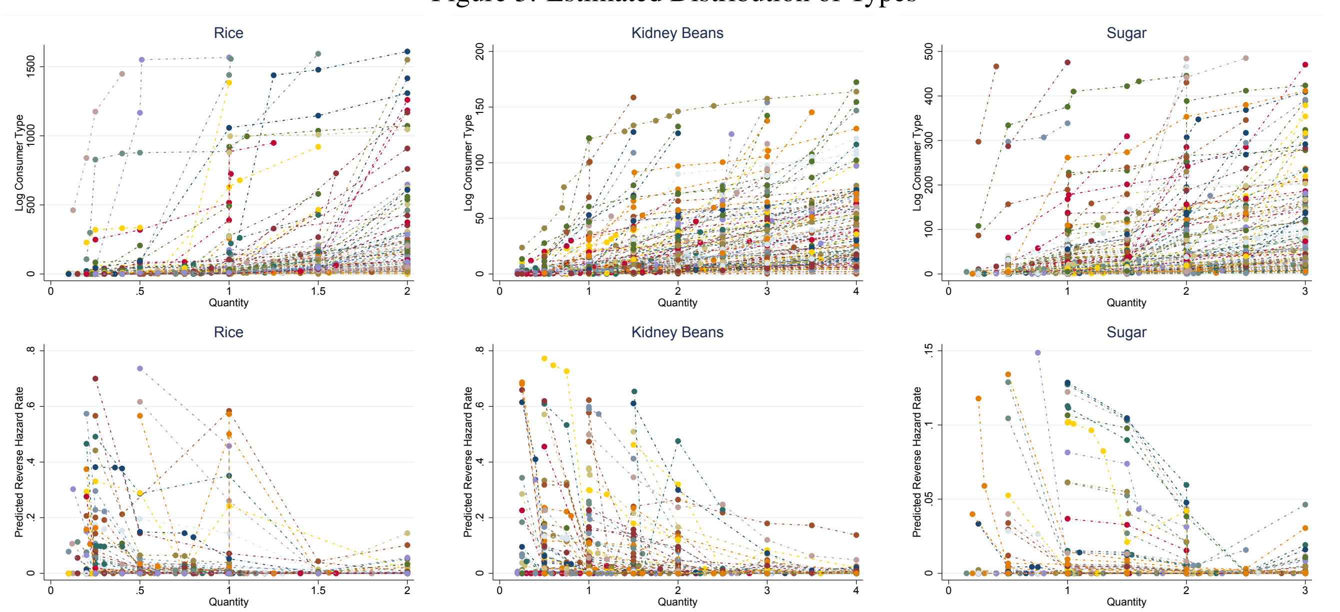
Figure 2: Estimated Multiplier on Participation (or Budget) Constraints



Estimates of Multipliers. Figure 2 reports the estimated multipliers on the participation (or budget) constraints in each village for each good under the quadratic specification of the multiplier in equation (16); see Appendix B for the linear specification. Note that, by construction, the estimated multiplier, $\gamma(\theta(q))$, ranges between 0 and 1 in each village for each good. We estimate that its mean across quantities and villages is 0.727 for rice with a standard deviation of 0.280, 0.713 for kidney beans with a standard deviation of 0.214, and 0.735 for sugar with a standard deviation of 0.227. For each good, the estimated multiplier is highly variable across quantities and clearly smaller than one for several of them: the 25th, 50th, and 75th percentiles of the distribution of $\gamma(\theta(q))$ across quantities and villages are 0.481, 0.875, and 0.963 for rice, 0.500, 0.750, and 0.899 for kidney beans, and 0.559, 0.810, and 0.919 for sugar. As discussed, the shape of the function $\gamma(\theta(q))$ discriminates across different instances of our model. In particular, the multiplier is estimated to be constant, as in the highly convex case of our model, only in a handful of villages. Furthermore, from tests of individual and joint significance of the parameters of $\gamma(\theta(q))$, we reject the hypothesis that the standard model applies ($\gamma(\theta(q))=1$ at all q) in all villages at standard significance levels.

For an intuition about why most villages do not conform to the standard model or, more generally, to the highly convex case of our model, recall that the seller's first-order condition can be expressed as in (15). For the highly convex case to apply, and therefore the multiplier to be constant, the term in brackets should replicate the variability of $G(q)$, since $x_3(q)$ is constrained to be positive and estimated to be approximately constant. Thus, $G(q)$ and $p(q)$ should be approximately inversely related. Since the unit prices $p(q)$ in the top panels of Figure 1 are decreasing and convex, $G(q)$ should then be increasing and concave. As Figure 1 shows, however, the curvature of $p(q)$ is most pronounced at *small* quantities, whereas the curvature of $G(q)$, which is initially convex, is most pronounced at *intermediate* quantities. This difference in the observed shapes of $p(q)$ and $G(q)$ must then be accommodated by $\gamma(\theta(q))$ varying across quantities, thus ruling out both the standard model and the highly convex case of our model.

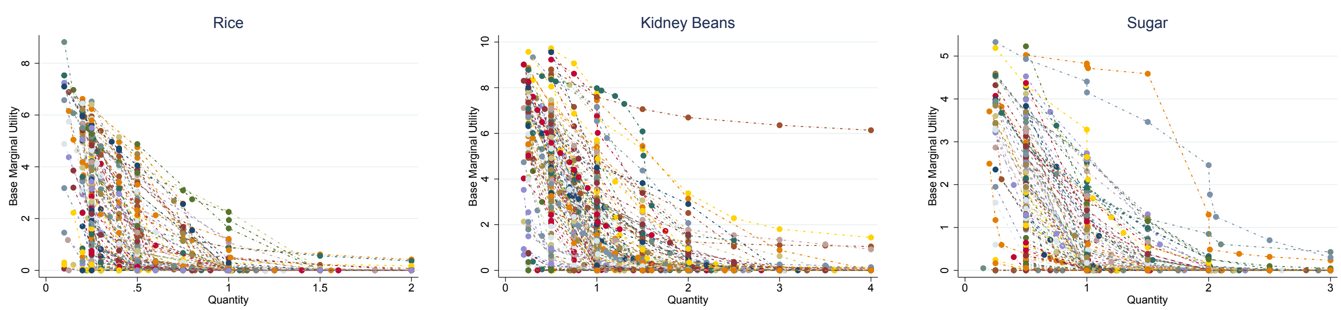
Figure 3: Estimated Distribution of Types



Estimates of Marginal Willingness to Pay. In the top panels of Figure 3, we plot our estimates of consumers' marginal willingness to pay as a function of quantity, $\theta(q)$, in each village for rice, kidney beans, and sugar. In the bottom panels of the same figure, we display our estimates of the reverse hazard rate of the distribution of consumer types in each village for each good as a function of quantity, $f(\theta(q))/F(\theta(q))$. (For readability, we trimmed the top 5% of estimates.) From the upward-sloping profiles in the top panels of Figure 3, we see that in each village, consumers' estimated marginal willingness to pay, θ, increases with quantity, as consistent with the incentive compatibility condition of our model. Also apparent from the bottom panels of the figure is that the estimated reverse hazard rate function, $f(\theta)/F(\theta)$, decreases nearly everywhere with θ and so with quantity, as is consistent with the monotone hazard rate condition sufficient for our assumption (PS). Since none of these restrictions have been imposed in estimation, we interpret these findings as validating our estimates of the type distribution. The t-statistics of these estimates reported in Appendix B further imply that these functions are estimated fairly precisely.

We note that the estimates of the support of types imply a much greater dispersion in consumers' marginal willingness to pay than in quantities purchased; recall from (12) that the level of θ is identified up to a normalization. Indeed, the bottom panels of Figure 1 show that most consumers purchase relatively small quantities: nearly all consumers consume less than 1 kilo of rice and less than 2 kilos of kidney beans and sugar per week. Instead, as the top panels of Figure 3 reveal, the support of types for each good is several times wider than the support of quantities. The fact that consumers markedly differ in their marginal willingness to pay for a good is important for the nonlinearity of observed prices, since it provides sellers with a strong incentive to discriminate across consumers. In particular, there is scope for sellers to distinguish consumers by their intensity of preference. We examine the extent to which sellers exert market power and extract surplus through price discrimination in Section 5.3.1.

Figure 4: Estimated Base Marginal Utility

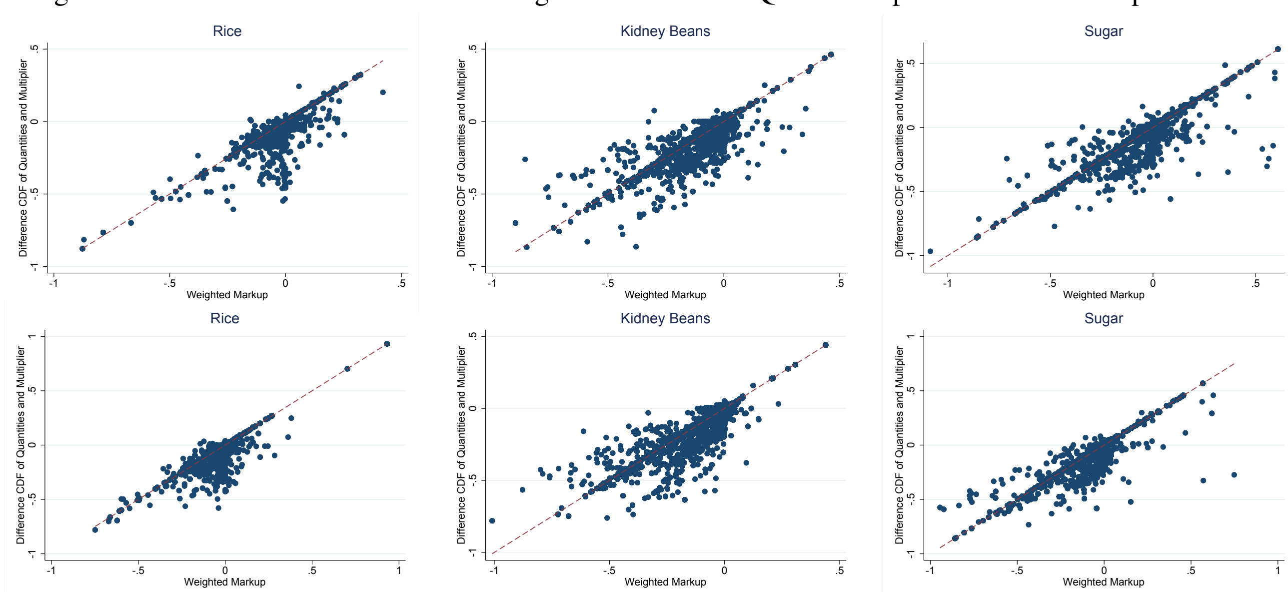


Estimates of Marginal Utility. In Figure 4, we plot the estimates of base marginal utility, $\nu'(q)$, at each quantity in each village and for each good. Note that $\nu'(q)$ decreases with quantity in all villages, as consistent with the model, even though no such monotonicity restriction has been imposed in estimation. Since consumers' marginal willingness to pay, $\theta(q)$, increases with quantity whereas their base marginal utility, $\nu'(q)$, decreases with quantity, marginal utility, $\theta(q)\nu'(q)$, decreases with q less rapidly than base marginal utility. The fact that marginal utility is nonetheless downward sloping implies that at the margin, each good is valued very differently at different levels of consumption. The large curvature in utility we estimate suggests the potential for rich distributional implications of nonlinear pricing, especially compared to alternative pricing schemes such as linear pricing, since different quantities are valued differently by consumers of any given taste. We further explore these implications in Section 5.3.1, where we link the *scope* for price discrimination, as captured by the distribution of consumers' tastes and marginal utility, to the *type* (first or second degree) of price discrimination that we infer sellers practice in our villages.

5.2.1 Model Fit

A key implication of our model is that the shape of the price schedule is determined by the cumulative distribution function of quantities, $G(q)$, and thus of consumers' marginal willingness to pay, $F(\theta)$, the associated probability density function, $f(\theta)$, and the multiplier on consumers' participation (or budget) constraints, $\gamma(\theta(q))$, as implied by a seller's first-order condition. Although the distribution of consumers' marginal willingness to pay and the multiplier are unobserved, they are directly related to the observed distributions of quantities and unit prices, as argued. Thus, one way to assess the fit of the model to the data is to determine the extent to which our estimates of the marginal cost $c'(Q)$, the multiplier $\gamma(\theta(q))$, and the auxiliary function $x_3(q)$ satisfy the relationship between the observed distribution of quantity, $G(q)$, and unit prices, $p(q)$, implied by a seller's first-order condition in (15) in each village for each quantity and good.

Figure 5: Model Fit within and across Villages for Linear and Quadratic Specification of Multiplier Function



To this purpose, we re-write equation (15) and plot in Figure 5 the estimated value of $G(q)-\gamma(\theta(q))$ on the y-axis against the markup measure $1/T'(q)-1/c'(Q)$ weighted by the auxiliary function $x_3(q)$, or *weighted markup* for brevity, on the x-axis, for each quantity and good in each village. To compute this markup measure, we use the fact that $T'(q)=t_1 p(q)$ by our specification of $T(q)$ in (13), with t_1 estimated and $p(q)$ observed. The closer the relationship between $G(q)-\gamma(\theta(q))$ and the weighted markup to the 45-degree line, the better the fit of the model to the data. In Figure 5, each dot represents the fit for a quantity of a given good in a given village: the top panels correspond to the case in which the multiplier $\gamma(\theta(q))$ is specified as a logistic function with a linear index in quantity, and the bottom panels correspond to the case in which the multiplier $\gamma(\theta(q))$ is specified as a logistic function with a quadratic index in quantity. The model fits the price and quantity data from each village remarkably well for each good. For instance, the R^2 of a linear regression line in each of the top panels is 0.953, 0.919, and 0.947 for rice, kidney beans, and sugar, respectively, whereas in each of the bottom panels, it is 0.968, 0.899, and 0.936. The R^2 for the bottom panels is lower partly because of the greater number of villages we successfully estimate in the quadratic case, as discussed in Section 5.1. We omit the corresponding regression lines from the figure since they are indistinguishable from the 45-degree lines.

5.3 Welfare Implications of the Estimated Model

5.3.1 Distortions Associated with Price Discrimination

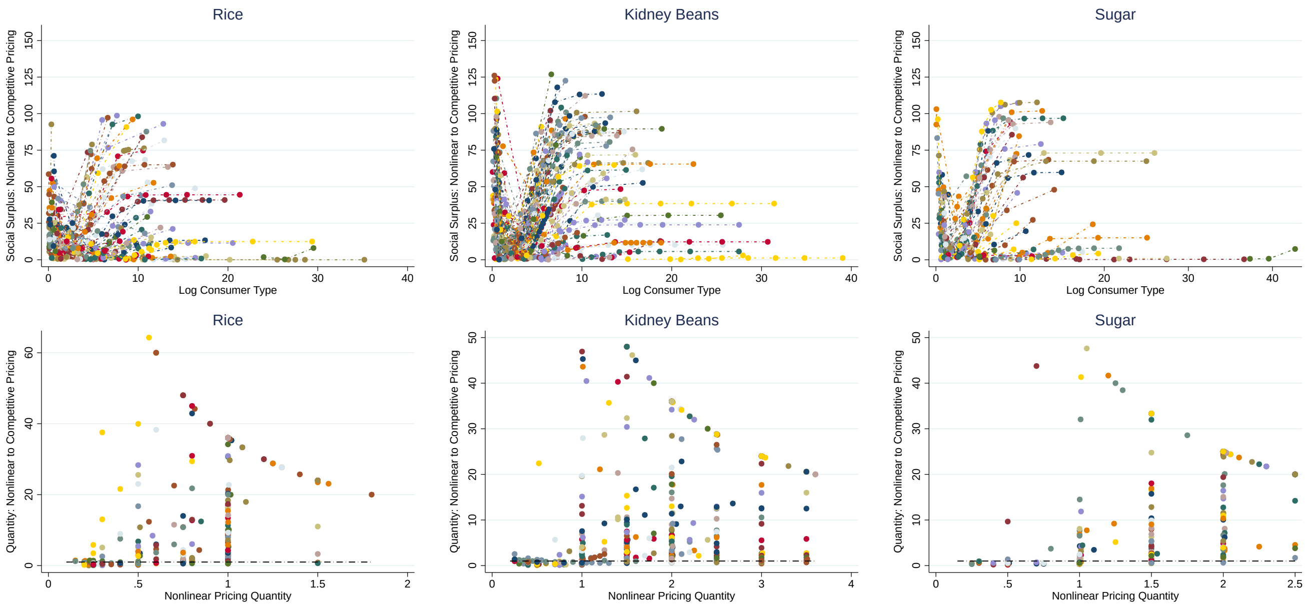
As discussed, our model allows for varying degrees of market power among sellers. Sellers' market power can distort the allocation of a good relative to first best, thereby reducing the gains from trade. It also affects the distribution of these gains between consumers and sellers. For instance, in the extreme case in which a seller could charge personalized prices and perfectly price discriminate, the resulting allocation would be efficient, but the seller would obtain all surplus. Alternatively, sellers could practice less efficient forms of price discrimination, of the second- or third-degree type, leading to allocations that do not maximize social surplus but in which larger surplus fractions accrue to consumers.

Here we examine the size of the distortions induced by sellers' market power in each village and their distributional implications, as implied by our estimates, by comparing social surplus and quantities consumed under observed nonlinear pricing and under first best. We interpret first best as a scenario in which free entry in a market is possible, so sellers price at cost. Since only the marginal cost of the total quantity of a good is identified in each village, for this exercise we assume the simplest possible (weakly convex) cost schedule for each village, whereby marginal cost is constant at all quantities—and so equal to average cost—except at the largest one, where it is given by our estimate of $c'(Q)$. Since a seller must break even at each quantity, we further assume that the average cost of any quantity different from the largest one equals the lowest observed unit price of these quantities. Hence, a seller's profit is nonnegative at each such quantity.

In the top panels of Figure 6, we graph the percentage gains in social surplus against (log) consumer types, when moving from observed nonlinear pricing to the counterfactual first-best scenario in which unit prices equal cost for each village, good, and quantity considered. We compute these gains as $\Delta SS_{fb} = [SS_{fb}(\theta)/SS_{np}(\theta) - 1]100$, where the subscript *fb* stands for first-best pricing and *np* for nonlinear pricing. The dotted lines join social surplus across (log) consumer types in each village. Note that the loss in social surplus implied by nonlinear pricing, across quantities and villages, ranges from about zero to over 100% of the surplus under nonlinear pricing, is especially large for kidney beans, and is approximately U-shaped over the interval where most types are. In particular, a sizable fraction of consumers of low to intermediate quantities do not suffer large distortions. (In the figure, we have trimmed the top 2% of such changes for readability, so the figure slightly underestimates actual losses from nonlinear pricing.)

In the bottom panels of Figure 6, we plot the ratio q_{fb}/q_{np} of first-best quantities, q_{fb}, to observed quantities, q_{np}, as a function of q_{np}. In most villages, the consumption of households with intermediate to large valuations of a good is most compressed relative to first best: the ratio is maximal at intermediate to large quantities. Interestingly, several consumers of the smallest quantities consume *above* first best; that is, the ratio q_{fb}/q_{np} is *smaller* than one for these households. Hence, sellers overall practice an inefficient form of price discrimination, which, however, leads consumers of the smallest quantities to overconsume rather than, as commonly argued, underconsume.

Figure 6: Social Surplus and Quantities under Nonlinear Pricing vs. First-Best Pricing



5.3.2 Nonlinear vs. Linear Pricing

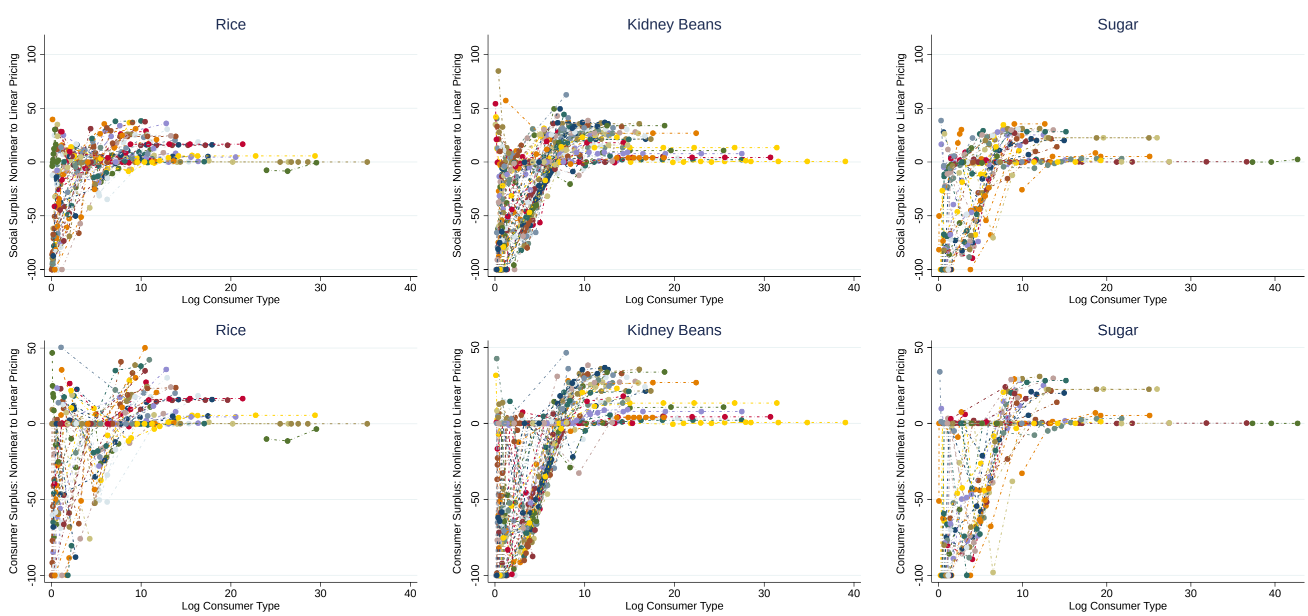
It has been argued that the ability of sellers to price discriminate through quantity discounts in developing countries hurts poor consumers more than rich consumers. In particular, quantity discounts may limit the access of the poorest households to basic goods and services, since these households tend to purchase the smallest quantities and thereby face the highest unit prices; see Attanasio and Frayne (2006) for references. Based on our model and estimates, we can examine which households are hurt (or benefit) more from the price discrimination we observe relative to linear pricing, as well as the efficiency of nonlinear and linear pricing, by comparing consumer and social surplus under nonlinear pricing and under the counterfactual scenario that would emerge if sellers were constrained to price linearly, for instance, by regulation. Importantly, this exercise entails not just a comparison of the price and quantity combinations under the two pricing schemes, but also of the *size* of the market served under each scheme. As formalized in Propositions 5 and 6, a seller who is prevented from discriminating may end up excluding some consumers under linear pricing. We find that this is precisely the case for many villages in our sample. We can also compare nonlinear to linear pricing in our model and in the standard model. Since the standard model is rejected in all villages, this exercise helps to shed light on the nature of the bias that would arise if the standard model was incorrectly presumed to apply.

Our Model. To compare the allocations observed under nonlinear pricing to those that would emerge counterfactually under linear pricing, we need an estimate of consumers' reservation utility to determine which households participate under linear pricing. As discussed in Section 4.1, though, reservation utility is only identified for consumer types whose participation (or budget) constraints bind.

In the absence of a point estimate, we proceed as follows. We set the reservation utility of the lowest type equal to this type's estimated utility, $\overline{u}(\underline{\theta}){=}u(\underline{\theta})$. For any other type, we set $\overline{u}(\theta){=}u(\theta)$ if the change in the estimated multiplier across the corresponding quantities is significantly different from zero (in which case we infer that the relevant constraint binds at θ) or we set $\overline{u}(\theta){=}\overline{u}(\theta')$ otherwise (in which case we infer that the relevant constraint does not bind at θ), where θ' is the largest estimated type smaller than θ whose constraint binds. When we cannot reject that the multiplier is constant at all quantities, and so the relevant constraint does not bind anywhere in the interior of the interval of types, we bound consumers' reservation utility from below and above. In particular, we bound reservation utility from *below* by setting $\overline{u}(\theta){=}u(\underline{\theta})$ except for the highest type, whose reservation utility we set at $u(\overline{\theta})$. We refer to this case as the *low reservation utility* case. Alternatively, we bound reservation utility from *above* by setting $\overline{u}(\theta){=}u(\theta)$ for each θ and refer to this case as the *high reservation utility* case, in which reservation utility is the highest possible one consistent with our model. See Appendix B for details.

Given these reservation utilities, we compute the percentage gains in consumer surplus, $\Delta CS_{lp}{=}[CS_{lp}(\theta)$-$/CS_{np}(\theta)-1]100$, and social surplus, $\Delta SS_{lp}=[SS_{lp}(\theta)/SS_{np}(\theta)-1]100$, when switching from nonlinear to linear pricing, where *lp* stands for linear pricing and *np* for nonlinear pricing. We then plot these gains (for each of the assumed reservation utility schedules) against (log) consumers types. Here we consider the low reservation utility case and discuss the high reservation utility case in Appendix B. In this latter case, results are qualitatively and quantitatively similar, except that a high reservation utility implies lower linear prices and so a somewhat higher consumer surplus under linear pricing than implied by the low reservation utility case. In both cases, consumers of small to intermediate quantities largely benefit from nonlinear pricing, that is, $\Delta CS_{lp}<0$, especially purchasers of the smallest quantities, and so prefer nonlinear to linear pricing.

Figure 7: Nonlinear vs. Linear Pricing under Augmented Model (Low Reservation Utility)



Consider then the low reservation utility case. In the top panels of Figure 7, we plot the social surplus gains from switching to linear pricing against (log) consumer types, and in the bottom panels, we plot the corresponding consumer surplus gains. As the top panels of Figure 7 show, nonlinear pricing leads to greater social surplus for most consumers with low to intermediate types, especially for rice and kidney beans.[29] Also, as the bottom panels show, a large fraction of consumers of small quantities are *better off* under nonlinear pricing, in particular for kidney beans—note the range of the y-axis is first negative then positive in all panels. That is, for a sizable fraction of consumers who purchase small quantities under nonlinear pricing, the greater consumption that nonlinear pricing gives rise to leads to higher levels of consumer and social surplus. On the contrary, consumers of intermediate to large quantities are better off under linear pricing—for them, consumer and social surplus are higher under linear pricing. Intuitively, under linear pricing, sellers provide smaller quantities, thereby inducing consumers to purchase less, but also charge lower prices. From the bottom panels of Figure 7, the benefit of lower prices indeed outweighs the utility loss from lower consumption only for intermediate to high types who consume intermediate to large quantities: their consumer surplus is higher under linear pricing.[30] The reduced ability of sellers to exert market power under linear pricing also implies a lower producer surplus from nearly all types than under nonlinear pricing.

A critical reason why consumer and social surplus are higher for some consumer types under nonlinear pricing is the higher degree of market participation that nonlinear pricing generates. In fact, since we observe quantity discounts, nonlinear pricing can increase consumer surplus only if linear pricing leads to the exclusion of some consumers by Propositions 5 and 6. We can measure the degree of exclusion that linear pricing induces by computing the fraction of consumers who would not participate in the market under linear pricing. The fractions of excluded consumers below the 25th, between the 25th and the 75th, and above the 75th percentile of types across villages are 0.635, 0.340, and 0.021 for rice; 0.839, 0.626, and 0.007 for kidney beans; and 0.802, 0.607, and 0.014 for sugar, respectively. Since nearly all consumers participate under observed nonlinear pricing, it follows that a large fraction of consumers, mostly purchasers of small quantities, would be excluded under linear pricing and thus benefit from nonlinear pricing.[31]

Standard Model. Nonlinear pricing has very different implications in the economy described by our model relative to that described by the standard model of Maskin and Riley (1984). To stress this point, we consider a scenario in which we counterfactually assume that the standard model applies to all villages. Formally, we assume that $\overline{u}(\theta) = u(\underline{\theta})$ so that $\gamma(\theta) = 1$ for all consumers and reestimate the model's primitives under this assumption in each village for each good. See Appendix B for the estimates of marginal cost, marginal

[29]Social surplus is virtually unchanged across nonlinear and linear pricing for consumers who do not participate under linear pricing, and thus experience utility $\overline{u}(\theta)$, but participate and obtain utility close to $\overline{u}(\theta)$ under nonlinear pricing. As apparent from Figure 7, this latter group of consumers are either the smallest or the largest types.

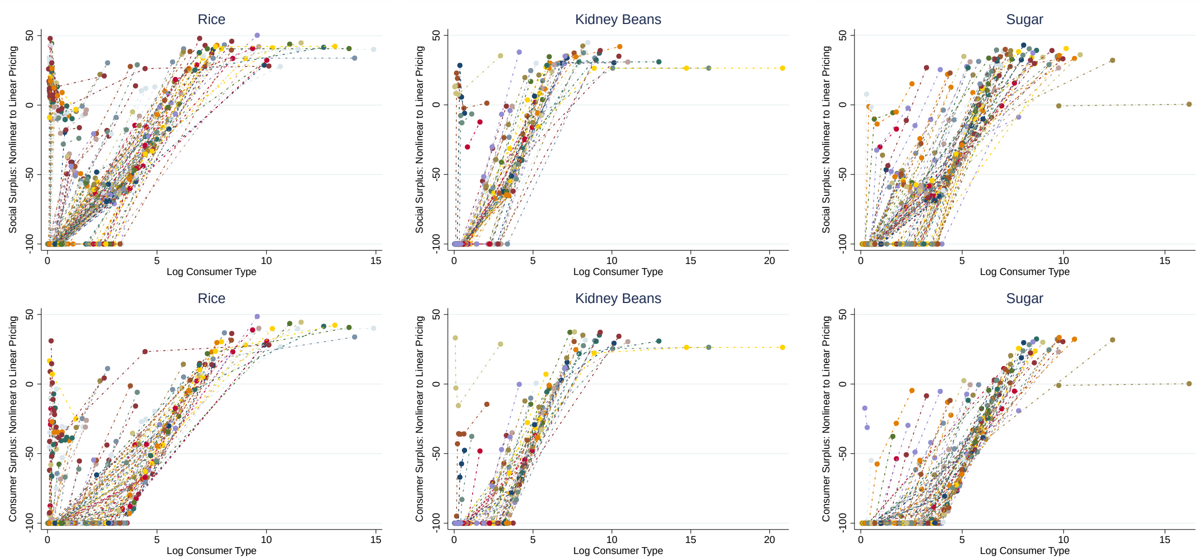
[30]Consumer surplus changes close to zero are those of types whose average price paid under nonlinear pricing is close to the linear price. Changes of consumer surplus by -100% or so are experienced by consumers with low taste for a good who do not participate under linear pricing. Their utility under linear pricing then equals their reservation utility, which is estimated at close to zero in many villages. An analogous argument holds for social surplus changes by -100% or so.

[31]The pattern of exclusion implied by linear pricing depends in general on the distribution of consumer types and the shape of reservation utility. Intuitively, though, since higher types are more profitable than lower ones, a seller typically faces a stronger incentive to exclude lower types than higher ones. Based on our estimates, high linear prices are optimal and lead consumers of relatively small quantities to be excluded. The reason is that the large taste parameters we estimate for households who purchase intermediate to large quantities imply that sellers can more than make up for excluding low types by charging high prices to the remaining ones.

willingness to pay, and base marginal utility, when the standard model is incorrectly assumed to apply. Given these new estimated primitives, we then compute consumer and social surplus under the observed nonlinear pricing allocation, under first best, and under linear pricing.

We find that, contrary to our model, the standard model implies that the greatest distortions, relative to first best, are suffered by the households who consume the smallest quantities and that these households consume quantities much smaller than under first best. Relative to our model, the standard model not only overestimates the loss in consumer and social surplus for consumers of small quantities, when moving from nonlinear to linear pricing, but also overestimates the corresponding gains for consumers of intermediate to large quantities. Indeed, by comparing the bottom panels of Figure 8 with those of Figure 7, it is apparent that consumers with low to intermediate types tend to benefit more from nonlinear pricing under the standard model than under our model whereas consumers with intermediate to high types tend to benefit more from linear pricing under the standard model than under our model. The top panels of Figures 7 and 8 also show that the standard model overestimates the loss in social surplus for consumers of small quantities, and the gains in social surplus for consumers of large quantities, when moving to linear pricing.

Figure 8: Nonlinear vs. Linear Pricing under Standard Model



To understand these findings, note that two countervailing forces are at play. On the one hand, the standard model requires that $[\theta\nu'(q(\theta)) - c'(Q)]/\nu'(q(\theta))$ be equal to $[1 - F(\theta)]/f(\theta)$, which decreases with θ by assumption (PS), as discussed. Hence, the standard model necessarily implies greater consumption distortions for lower consumer types than for higher ones relative to first best, contrary to what we estimate. This monotonicity of consumption distortions, in turn, tends to make nonlinear pricing less desirable for consumers of small quantities with low marginal willingness to pay but more desirable for consumers of large quantities with intermediate to high marginal willingness to pay. On the other hand, since the reservation utility profile is flat in the standard model, sellers have a greater ability to extract consumer surplus through linear pricing under the standard model than under our model. Sellers may thus induce greater consumer exclusion under linear pricing than under our model, even among purchasers of large quantities. On balance, this second force dominates in many villages for low consumer types: many more low and intermediate consumer types are excluded under linear pricing according to the standard model than according to our model, as apparent by comparing the bottom panels of Figures 7 and 8: much more mass is piled at -100%, which signifies consumer exclusion, in the panels for the standard model than in those for our model.

Discussion. Even small deviations of $\gamma(\theta(q))$ from one—its value in the standard model—which we estimate for several quantities and villages, are associated with very different behavior on the part of sellers and consumers. These deviations have important distributional implications and thus are key to evaluating alternative pricing mechanisms. In particular, assuming counterfactually that the standard model applies, when it is rejected instead, can lead to incorrect inferences about the impact of nonlinear pricing and the relative benefits of nonlinear and linear pricing.

5.4 The Effect of Income Transfers

As discussed in Section 3.3, the version of our model with budget-constrained consumers can be used to evaluate the impact of a targeted transfer, such as the Progresa one, on consumption and prices. Here we do so and find that the Progresa program had a significant impact on prices. Note that by Proposition 7, income transfers not only encourage greater consumption but also induce sellers to modify their price schedules in response to consumers' greater ability to pay, typically by charging higher prices to some (if not all) consumers. Indeed, it has been documented that food expenditure per adult equivalent has increased by 13% among eligible households as a result of Progresa; see, for instance, Angelucci and De Giorgi (2009). A small literature has also examined the effect of Progresa on the prices of agricultural commodities. As mentioned, however, Hodinott et al. (2000) and Angelucci and De Giorgi (2009) found no evidence that the Progresa transfer has induced a systematic increase in the (average unit) price of basic staples.

In this section, we assess the extent to which the Progresa transfer has affected prices, in particular, the degree of price discrimination, and consumer surplus. Unlike the studies just mentioned that focus only on the impact of transfers on *average* unit prices, we examine the impact of Progresa on their entire schedule. We find that Progresa has had a significant effect on prices but that this effect *cannot* be detected without accounting for the nonlinearity of unit prices. Moreover, we show that our model can account for the shift in the price schedule induced by the program.

Transfers and Prices. We measure the impact of the Progresa transfer on prices based on a second-order bivariate Taylor expansion in $\log(q)$ and $[1 - G(q)]/g(q)$ of a seller's first-order condition,

$$\log[p(q)] \approx \beta_0 + \beta_1 \log(q) + \beta_2 \left[\frac{1 - G(q)}{g(q)}\right] + \beta_3 \log(q) \left[\frac{1 - G(q)}{g(q)}\right] + \beta_4 \log(q)^2 + \beta_5 \left[\frac{1 - G(q)}{g(q)}\right]^2, \quad (18)$$

derived at the end of Appendix A. In this expansion, the multiplier $\gamma(\theta(q))$ is treated as a function of quantity. This reduced form relates log unit prices, $\log[p(q)]$, to log quantities, $\log(q)$, and the inverse hazard rate of the distribution of quantities, $[1-G(q)]/g(q)$, in each village. Note that this latter term captures the importance of the shape of the distribution of consumers' marginal willingness to pay for prices, as consistent with our model. Unit prices are related to the *distribution* of consumer preferences in a market, in particular to its inverse hazard rate, and so to the distribution of quantities because of the one-to-one relationship between consumer tastes and demand from $F(\theta) = G(q)$ and $q = q(\theta)$.

To start, in the first column of Table 2, we estimate the effect of Progresa on the average prices of rice, kidney beans, and sugar in the localities in our sample. Specifically, we regress observed unit prices on a constant and an indicator of the program, the dummy variable "Treatment," which equals one in (randomly) treated villages; we include wave fixed effects and cluster standard errors at the village level. Consistently with the existing literature on the price effects of Progresa, the Progresa transfer effectively has had no impact on average unit prices: the estimated coefficient on the "Treatment" dummy for each good is small and not significantly different from zero.

Our model implies that the effect of an income transfer on unit prices, $p(q) = T(q)/q$, is ambiguous, even when it may give rise to substantial changes in the overall price schedule, since households' expenditure, $T(q)$, and consumption, q, may both increase. But Corollary 2 also implies that even no increase or a modest increase in average unit prices after a transfer can be associated with a substantial change in the price schedule, leading to a greater intensity of price discrimination overall, which we observe in our data. Therefore, we next examine how the Progresa transfer has affected the magnitude of quantity discounts. To this purpose, in the second column of Table 2, we report the results of a regression of log unit prices on the "Treatment" dummy, log quantity, and an interaction between these two variables. Note that this is an exercise analogous to that of Table 1, except that we allow the Progresa transfer to shift both the intercept and the slope of the price schedule. For each commodity, we find that the coefficient on log quantity becomes more negative, and significantly so for kidney beans and sugar, that is, the intensity of price discrimination increases with the Progresa transfer. The effect of log quantity on log unit prices increases in absolute value from -0.212 to -0.256 for rice, from -0.114 to -0.154 for kidney beans, and from -0.123 to -0.185 for sugar.[32]

We then assess the extent to which our model accounts for the observed nonlinearity of prices as well as the change in the unit price schedule resulting from Progresa. To this end, we first estimate (18) and next a version of it augmented to account for the impact of the Progresa transfer through the "Treatment" dummy and the interaction between this dummy and log quantity. We report the results of this estimation in columns 3 and 4 of Table 2. We stress that the inverse hazard rate of quantities, $[1 - G(q)]/g(q)$, in both regressions is computed for *each village*.[33] This approach, therefore, allows for heterogeneous impacts of the program across villages and, since the program obviously affected the inverse hazard function of quantities in

[32]Although this regression captures the variability of unit prices both across quantities within a village and across villages, the aggregation implicit in this exercise is valid. In particular, we could allow for village fixed effects to capture the unobserved variability in marginal costs across villages. We notice that both village fixed effects and the coefficients on the inverse hazards are identified, since we have several waves of data and the inverse hazards vary also across quantities. The results we obtain with or without village fixed effects are very similar, though, consistently with the fact that marginal costs estimated at the locality level do not display a large degree of variation across localities; see Appendix B.3.3. Effectively, the variation in the inverse hazards captures most of the variation in price schedules. Therefore, we present here the results of the simpler estimation without fixed effects and illustrate the fixed effect estimates in Appendix C.

[33]Standard errors are computed by bootstrap to account for the fact that inverse hazard rates are estimated too.

"treated" villages, could be interpreted as a *mediation analysis* of the effect of Progresa on the price schedule.

In column 3, where we report the estimates of (18), we notice that both the inverse hazard rate, $[1 - G(q)]/g(q)$, and its interaction with $\log(q)$ are significant for each of the three commodities—strongly so for rice and kidney beans. The quadratic inverse hazard rate is not significant for any commodity, whereas the quadratic term in $\log(q)$ is significant for kidney beans and strongly so for sugar.

In column 4, we report the estimates of the augmented version of equation (18) that accounts for the Progresa transfer through the "Treatment" dummy and the interaction between this dummy and $\log(q)$. Importantly, we see that the "Treatment" dummy is not significant and, unlike in column 2 for kidney beans and sugar, does not significantly affect the impact of $\log(q)$ on log unit prices. We also notice that the point estimates of the coefficient on $\log(q)$ interacted with the "Treatment" dummy are greatly reduced in absolute value. These results thus indicate that our model is capable of explaining the shift in price schedules documented in column 2. Specifically, a significant portion of the change in prices induced by the program is accounted for by the change in the distribution of quantities consumed in each village, in particular in the curvature of this distribution as captured by its inverse hazard rate, rather than by the direct effect of the transfer on the intercept or slope of the price schedule. Indeed, unlike the "Treatment" dummy and its interaction with $\log(q)$, the coefficients on $\log(q)$, $[1 - G(q)]/g(q)$ and their interaction are all significant in column 4 of Table 2.

Price Changes and Consumer Surplus. The reduced-form analysis just presented shows that our model can explain the substantial shifts in price schedules induced by the Progresa transfer. This evidence and the mediation analysis provided in Table 2 support the budget constraint interpretation of our model. In light of this evidence, we can use our model, in which consumers' preferences and subsistence or budget constraints are explicitly formalized, to further assess the impact of Progresa on consumer surplus. Specifically, given our estimates of consumer surplus in the localities of our sample, we can compute the elasticity of consumer surplus to unit prices, which ranges across villages between -91.91% and -0.33% with a median (mean) of -7.50% (-10.96%) for rice, between -89.04% and 0.00% with a median (mean) of -9.86% (-11.23%) for kidney beans, and between -81.08% and 0.00% with a median (mean) of -8.51% (-10.14%) for sugar in the low reservation utility case. The elasticity of consumer surplus to quantity consumed is much smaller, ranging from 0.10% to 27.40% for rice, from 0.00% to 24.56% for kidney beans, and from 0.00% to 22.58% for sugar. Hence, ignoring the price change associated with the Progresa transfer can substantially overstate the increase in consumer surplus from the program: in some villages, the price increase almost erases households' gains from consuming any of the goods before the transfer took place.

Discussion. These results are important in assessing the impact of cash transfers, a common policy tool in developing countries. Cash transfers may not affect, on average, unit prices in a market but may lead to an overall increase in the intensity of price discrimination, as we observe in our data. Such a change in prices has an impact not just on households that are beneficiaries of the program but also on non-eligible households. Since all households are affected by the resulting price change, cash transfers may then have a more limited beneficial effect than is commonly estimated, because of their partially positive *direct effect* on

Table 2: Effect of Cash Transfers on Prices

	Rice Unit Values				Kidney Beans Unit Values				Sugar Unit Values			
	1	2	3	4	1	2	3	4	1	2	3	4
Intercept	1.947***	1.868***	1.864***	1.869***	2.375***	2.420***	2.421***	2.423***	1.743***	1.775***	1.777***	1.769***
	(0.0105)	(0.0127)	(0.00711)	(0.0121)	(0.00940)	(0.00952)	(0.00767)	(0.00998)	(0.00403)	(0.00660)	(0.00448)	(0.00618)
Treatment	−0.00750	−0.0123		−0.00663	−0.0142	0.00600		−0.00494	0.00556	0.0265**		0.0120
	(0.0112)	(0.0149)		(0.0142)	(0.0117)	(0.0127)		(0.0112)	(0.00485)	(0.00909)		(0.00660)
$\log(q)$		−0.212***	−0.181***	−0.161***		−0.114***	−0.172***	−0.160***		−0.123***	−0.182***	−0.168***
		(0.0175)	(0.0324)	(0.0413)		(0.0109)	(0.0281)	(0.0294)		(0.0183)	(0.0172)	(0.0220)
$\frac{1-G(q)}{g(q)}$			−0.0110**	−0.0111**			−0.00505***	−0.00503***			−0.00275*	−0.00273*
			(0.00423)	(0.00422)			(0.00110)	(0.00110)			(0.00107)	(0.00107)
$\log(q) \times \frac{1-G(q)}{g(q)}$			−0.00951*	−0.00943*			−0.00555*	−0.00554*			−0.00546*	−0.00543*
			(0.00469)	(0.00469)			(0.00219)	(0.00219)			(0.00219)	(0.00218)
$\log(q)$ × Treatment		−0.0441		−0.0270		−0.0399*		−0.0152		−0.0617*		−0.0173
		(0.0255)		(0.0228)		(0.0191)		(0.0136)		(0.0271)		(0.0188)
$\log(q)^2$			0.0458	0.0457			0.0526*	0.0523*			0.0813***	0.0810***
			(0.0398)	(0.0395)			(0.0251)	(0.0250)			(0.0174)	(0.0173)
$\left[\frac{1-G(q)}{g(q)}\right]^2$			0.0000243	0.0000261			0.00000484	0.00000473			−0.0000173	−0.0000171
			(0.0000469)	(0.0000470)			(0.0000213)	(0.0000212)			(0.0000242)	(0.0000242)
1999-March	0.0500***	0.0356***	0.0355***	0.0362***	−0.0827***	−0.0872***	−0.0867***	−0.0865***	0.00888	0.0118*	0.0115*	0.0117*
	(0.00858)	(0.00757)	(0.00788)	(0.00777)	(0.00785)	(0.00743)	(0.00744)	(0.00743)	(0.00551)	(0.00528)	(0.00518)	(0.00516)
1999-Nov	0.0501***	0.0355***	0.0357***	0.0362***	−0.115***	−0.114***	−0.111***	−0.111***	0.0549***	0.0583***	0.0592***	0.0594***
	(0.00772)	(0.00744)	(0.00749)	(0.00739)	(0.00734)	(0.00722)	(0.00719)	(0.00720)	(0.00463)	(0.00439)	(0.00429)	(0.00427)
2000-Nov	0.0302***	0.0213**	0.0229***	0.0217**	−0.196***	−0.203***	−0.205***	−0.201***	0.0307***	0.0283***	0.0330***	0.0303***
	(0.00793)	(0.00679)	(0.00662)	(0.00767)	(0.0122)	(0.0115)	(0.0109)	(0.0114)	(0.00581)	(0.00567)	(0.00518)	(0.00565)
2003	−0.0146	0.00932	0.00536	0.00504	−0.0207	−0.0316**	−0.0348***	−0.0307**	0.170***	0.174***	0.172***	0.169***
	(0.0101)	(0.00996)	(0.00924)	(0.00954)	(0.0110)	(0.0101)	(0.00894)	(0.0101)	(0.00724)	(0.00693)	(0.00611)	(0.00654)
R^2 adj.	0.00608	0.140	0.165	0.166	0.0442	0.0946	0.130	0.130	0.0468	0.118	0.201	0.201
Observations	62,368	62,368	62,368	62,368	82,024	82,024	82,024	82,024	91,782	91,782	91,782	91,782

Note: the superscript * stands for $p < 0.05$, the superscript ** for $p < 0.01$, and the superscript *** for $p < 0.001$. Standard errors are clustered at the locality level.

eligible households and their negative *indirect effect* on non-eligible ones.[34]

6 Conclusion

We propose a model of nonlinear pricing in which consumers differ in their tastes for goods, face heterogeneous subsistence constraints leading to heterogeneous budget constraints for a seller's good, and have access to different outside options to participating in a market. Incorporating budget constraints is an important advance in this literature, since it makes the model relevant for several contexts of practical relevance. The model we propose also encompasses a wide variety of market structures, ranging from pure monopoly to competition. In all these settings, the implications of nonlinear pricing for consumer, producer, and social surplus are fundamentally different from those arising from standard models of nonlinear pricing, in which outside options are identical across consumers and consumers are assumed unconstrained in their purchase decisions. In particular, quantity discounts for *large* volumes can be associated with consumption above first best at *low* volumes.

We prove that this more general model is identified under common assumptions from information on prices and quantities purchased. We derive estimators of the model's primitives that can be readily implemented using a variety of publicly available data sets. We use the publicly available data on the evaluation of a large and celebrated conditional cash transfer program, Progresa, to estimate the model, which fits the data extremely well. Our empirical results have important implications for the relative desirability of nonlinear and linear pricing. In particular, we find that most consumers of small to intermediate quantities, typically the poorest ones, benefit from nonlinear pricing, despite the fact that sellers price discriminate through distortionary quantity discounts. We also find that nonlinear pricing leads to a greater degree of market participation, especially for consumers of the smallest quantities, which is all the more critical for the marginalized villages in our data in which the consumption of several households is at levels of subsistence.

Crucially, though, we show that by increasing consumers' ability to pay, cash transfers provide sellers with the opportunity to extract more surplus through nonlinear pricing. Thus, cash transfers in general lead to an increase in the intensity of price discrimination, as we document in the case of Progresa. Indeed, our paper is one of the first to uncover important shifts in the price schedule in the villages included in the Progresa evaluation sample. Specifically, we find that cash transfers implemented by Progresa have had a significant impact on prices in our villages, unlike what is commonly found in the literature. Furthermore, we show that our model can explain the observed shifts in price schedules.

This result is all the more relevant since cash transfers have become an increasingly popular policy tool in Latin America and many other developing countries. A few studies have analyzed the effect of transfers on the price of commodities and the consensus so far seems to have been that Progresa did *not* have appreciable effects on local prices. We prove that if we look beyond average prices, this conclusion is unwarranted. When the dependence of unit prices on quantity is taken into account, the price effect is substantial, in particular, the program is associated with an increase in the degree of price discrimination. Such a change

[34] This argument, though, neglects the positive spillovers on non-eligible households found by Angelucci and De Giorgi (2009).

in the equilibrium price schedule has had an impact not just on households beneficiaries of the program but also on non-eligible households, since all households have been affected by the overall price change. Cash transfers can then be less beneficial than typically inferred.

Overall, our estimation results thus suggest the importance of accounting for heterogeneity in consumers' preferences, constraints, and consumption opportunities when assessing the impact of nonlinear pricing.

References

ANGELUCCI, M., and G. DE GIORGI (2009): "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?," *American Economic Review*, 99(1), 486-508.

ATTANASIO, O., V. DI MARO, V. LECHENE, and D. PHILLIPS (2013): "Welfare Consequences of Food Prices Increases: Evidence from Rural Mexico," *Journal of Development Economics*, 104(C), 136-151.

ATTANASIO, O., and C. FRAYNE (2006): "Do the Poor Pay More?," mimeo.

ATTANASIO, O., AND E. PASTORINO (2010): "Nonlinear Pricing of Food in Village Economies: the Case of Mexico," mimeo.

BONTEMPS, C., and D. MARTIMORT (2014): Identification and Estimation of Incentive Contracts under Asymmetric Information: An Application to the French Water Sector," mimeo.

CAMPO, S., E. GUERRE, I. PERRIGNE, and Q. VUONG (2011): "Semiparametric Estimation of First-Price Auctions with Risk-Averse Bidders," *Review of Economic Studies*, 78(1), 112-147.

CHE, Y.-K., and I. GALE (2000): "The Optimal Mechanism for Selling to a Budget-Constrained Buyer," *Journal of Economic Theory*, 92(2), 198-233.

CRAWFORD, G.S., and M. SHUM (2007): "Monopoly Quality Degradation and Regulation in Cable Television," *Journal of Law and Economics*, 50(1), 181-209.

CRAWFORD, I., F. LAISNEY, and I. PRESTON (2003): "Estimation of Household Demand Systems with Theoretically Compatible Engel Curves and Unit Value Specifications," *Journal of Econometrics*, 114(2), 221-241.

CUNHA, J., G. DE GIORGI, and S. JAYACHANDRAN (2017): "The Price Effects of Cash Versus In-Kind Transfers," mimeo.

DEATON, A., (1989): "Rice Prices and Income Distribution in Thailand: A Non-parametric Analysis," *Economic Journal*, 99(395), 1-37.

EKELAND, I., J.J. HECKMAN, and L. NESHEIM (2004): "Identification and Estimation of Hedonic Models," *Journal of Political Economy*, 112(S1), S60-S109.

FABRICANT, S.J., C.W. KAMARA, and A. MILLS (1999): "*Why* the Poor Pay More: Household Curative Expenditures in Rural Sierra Leone," *International Journal of Health Planning and Management*, 14(3), 179-199; erratum in *International Journal of Health Planning and Management*, 14(4), 339-340.

GUERRE, E., I. PERRIGNE, and Q. VUONG (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525-574.

HECKMAN, J.J., R.L. MATZKIN, and L. NESHEIM (2010): "Nonparametric Identification and Estimation of Nonadditive Hedonic Models," *Econometrica*, 78(5), 1569-1591.

HODDINOTT, J., E. SKOUFIAS, and R. WASHBURN (2000): "The Impact of Progresa on Consumption: A Final Report," International Food Policy Research Institute, available at https://www.ifpri.org/publication/-impact-progresa-consumption.

JENSEN, R.T., and N.H. MILLER (2008): "Giffen Behavior and Subsistence Consumption," *American Economic Review*, 98(4), 1553-1577.

JULLIEN, B. (2000): "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93(1), 1-47.

KAHN, M.E., and F.A. WOLAK (2014): "Using Information to Improve the Effectiveness of Nonlinear Pricing: Evidence from a Field Experiment," mimeo.
LANCASTER, K.J. (1966): "A New Approach to Consumer Theory," *Journal of Political Economy*, 74(2), 132-157.
LIGON, E., J.P. THOMAS, and T. WORRALL (2002): "Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies," *Review of Economic Studies*, 69(1), 209-244.
LUO, Y., I. PERRIGNE, and Q. VUONG (2010): "Structural Analysis of Nonlinear Pricing," mimeo.
MASKIN, E., and J. RILEY (1984): "Monopoly with Incomplete Information," *RAND Journal of Economics*, 15(2), 171-196.
MCINTOSH, A.C. (2003): "The Poor Pay More for Their Water," *Habitat Debate*, 9(3), available at http://-www.unhabitat.org/hd/hdv9n3/12.asp.
MIRAVETE, E., and L.H. RÖLLER (2004): "Competitive Nonlinear Pricing in Duopoly Equilibrium: The Early U.S. Cellular Telephone Industry," mimeo.
PANNARUNOTHAI, S., and A. MILLS (1997): "The Poor Pay More: Health-Related Inequality in Thailand," *Social Science and Medicine*, 44(12), 1781-1790.
PERRIGNE, I., and Q. VUONG (2010): "Nonlinear Pricing in Yellow Pages," mimeo.
SEGAL, I., and S. TADELIS (2005): "Lectures in Contract Theory," mimeo.
STOLE, L.A. (2007): "Price Discrimination and Competition," in: *Handbook of Industrial Organization*, Vol. 3, M. Armstrong and R. Porter (eds.), Elsevier.
THOMAS, L. (2002): "Non-linear Pricing with Budget Constraint," *Economics Letters*, 75(2), 257-263.
WOLAK, F.A. (2014): "Designing Minimum-Risk Nonlinear Price Schedules for Water Utilities," mimeo.

A Model: Omitted Proofs and Details

Cumulative Multiplier in the Highly Convex Case: We derive here an expression that defines the multiplier in the highly convex case in terms of primitives. For simplicity, we consider the case in which $v(\theta, q) = \theta\nu(q)$ and $c'(q) = c$; the argument extends naturally to the more general case. Recall that in the highly convex case, the cumulative multiplier $\gamma(\theta)$ is equal to a constant, γ, at all points $\theta \in [\underline{\theta}, \overline{\theta})$. Here we solve for γ in the interesting case in which $0 < \gamma < 1$ for $[\underline{\theta}, \overline{\theta})$. In the remaining cases, the multiplier is trivial: $\gamma(\theta) = 1$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$ and $\gamma(\theta) = 0$ for all $\theta \in [\underline{\theta}, \overline{\theta})$. First, observe that (1) implies that

$$\nu'(q(\theta)) = cf(\theta)/[\theta f(\theta) + F(\theta) - \gamma], \tag{19}$$

so $q(\theta) = (\nu')^{-1}(cf(\theta)/[\theta f(\theta) + F(\theta) - \gamma])$. Recall that $u(\underline{\theta}) = \overline{u}(\underline{\theta})$ and $u(\overline{\theta}) = \overline{u}(\overline{\theta})$ when $0 < \gamma < 1$. Hence, $\overline{u}(\overline{\theta}) - \overline{u}(\underline{\theta}) = u(\overline{\theta}) - u(\underline{\theta})$ so γ is implicitly defined by

$$\overline{u}(\overline{\theta}) - \overline{u}(\underline{\theta}) = \int_{\underline{\theta}}^{\overline{\theta}} u'(x)dx = \int_{\underline{\theta}}^{\overline{\theta}} \nu(q(x))dx = \int_{\underline{\theta}}^{\overline{\theta}} \nu\left((\nu')^{-1}\left(\frac{cf(x)}{xf(x) + F(x) - \gamma}\right)\right)dx,$$

where the first equality follows from $\overline{u}(\overline{\theta}) - \overline{u}(\underline{\theta}) = u(\overline{\theta}) - u(\underline{\theta})$ and the fundamental theorem of calculus, and the second equality from the local incentive compatibility condition $u'(\theta) = \nu(q(\theta))$. $\square$

Proof of Proposition 1: Before proving Proposition 1, we first derive the simple BC problem and then establish that the first-order and complementary slackness conditions of the simple BC problem in (6) are necessary and sufficient to characterize an optimal menu. The proof of this result requires that assumptions analogous to those of potential separation, homogeneity, and full participation in the IR model hold in the BC model. We have discussed assumption (BCH) in the text and assumption (FP) in footnote (15), so we discuss here only assumption (FP). As in the IR model, the *potential separation* assumption in the BC model requires $l(\Phi, \theta)$ to be a weakly increasing function of θ for all $\Phi \in [0, 1]$, for which sufficient conditions are

$$\frac{\partial}{\partial\theta}\left(\frac{s_q(\theta, q)}{v_{\theta q}(\theta, q)}\right) \geq 0 \text{ and } \frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right) \geq 0 \geq \frac{d}{d\theta}\left(\frac{1 - F(\theta)}{f(\theta)}\right). \tag{20}$$

As explained in Jullien (2000), the first inequality in (20) implies that the conflict between rent extraction and efficiency is not too severe so that the marginal benefit of increasing the slope of the utility profile is weakly increasing with the type. When this occurs, a seller tends to desire convex quantity profiles, which implies that the monotonicity condition for $q(\theta)$ for incentive compatibility is easier to satisfy. The second and third inequalities in (20) amount to a simple strengthening of the monotone hazard rate condition ubiquitous in the mechanism design literature: as the type increases, the relative weight of types above θ compared with below θ decreases, and the seller becomes progressively more concerned about the "informational rents" left below θ.

To derive the simple BC problem in (6), we proceed in analogy with the derivation of the simple IR problem in the Supplementary Appendix. First, we rewrite the BC constraint as

$$I(\theta, q(\theta)) \geq t(\theta) = v(\theta, q(\theta)) - u(\theta), \tag{21}$$

since $u(\theta) = v(\theta, q(\theta)) - t(\theta)$. We presume $\overline{u}$ is low enough and then show that under the conditions of Proposition 1, (IR') is indeed redundant. The BC problem can be expressed in Lagrangian-type form as

$$\max_{\{u(\theta)\},\{q(\theta)\}\in\widehat{Q}} \left(\int_{\underline{\theta}}^{\overline{\theta}}[v(\theta, q(\theta)) - c(q(\theta)) - u(\theta)]f(\theta)d\theta + \int_{\underline{\theta}}^{\overline{\theta}}\{I(\theta, q(\theta)) - [v(\theta, q(\theta)) - u(\theta)]\}d\Phi(\theta)\right) \tag{22}$$

$$\text{s.t. } u'(\theta) = v_\theta(\theta, q(\theta)), \tag{23}$$

where $\widehat{Q}$ is the set of weakly increasing functions $q(\theta)$, and $\Phi(\theta)$ is the cumulative Lagrange multiplier on the budget constraints expressed as in (21). Next, note that

$$\int_{\underline{\theta}}^{\overline{\theta}} u(\theta)f(\theta)d\theta = u(\underline{\theta})\int_{\underline{\theta}}^{\overline{\theta}} f(\theta)d\theta + \int_{\underline{\theta}}^{\overline{\theta}}[u(\theta)-u(\underline{\theta})]f(\theta)d\theta = u(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}}\left(\int_{\underline{\theta}}^{\theta} u'(x)dx\right)f(\theta)d\theta.$$

Using the local incentive compatibility condition $u'(\theta) = v_\theta(\theta, q(\theta))$ and integrating by parts thus gives

$$\int_{\underline{\theta}}^{\overline{\theta}} u(\theta)f(\theta)d\theta = u(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}}\left(\int_{\underline{\theta}}^{\theta} v_\theta(x,q(x))dx\right)f(\theta)d\theta = u(\underline{\theta}) + \left(\int_{\underline{\theta}}^{\theta} v_\theta(x,q(x))dx\right)F(\theta)\Big|_{\underline{\theta}}^{\overline{\theta}}$$

$$-\int_{\underline{\theta}}^{\overline{\theta}} v_\theta(\theta,q(\theta))F(\theta)d\theta = u(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}} v_\theta(\theta,q(\theta))d\theta - \int_{\underline{\theta}}^{\overline{\theta}} v_\theta(\theta,q(\theta))F(\theta)d\theta. \tag{24}$$

Similarly,

$$\int_{\underline{\theta}}^{\overline{\theta}} u(\theta)d\Phi(\theta) = u(\underline{\theta})[\Phi(\overline{\theta})-\Phi(\underline{\theta})] + \left(\int_{\underline{\theta}}^{\theta} v_\theta(x,q(x))dx\right)\Phi(\theta)\Big|_{\underline{\theta}}^{\overline{\theta}}$$

$$-\int_{\underline{\theta}}^{\overline{\theta}} v_\theta(\theta,q(\theta))\Phi(\theta)d\theta = u(\underline{\theta})[\Phi(\overline{\theta})-\Phi(\underline{\theta})] + \Phi(\overline{\theta})\int_{\underline{\theta}}^{\overline{\theta}} v_\theta(\theta,q(\theta))d\theta - \int_{\underline{\theta}}^{\overline{\theta}} v_\theta(\theta,q(\theta))\Phi(\theta)d\theta. \tag{25}$$

Substituting (24) and (25) into the objective function in (22) yields

$$\int_{\underline{\theta}}^{\overline{\theta}}[v(\theta,q(\theta)) - c(q(\theta)) - u(\theta)]\,f(\theta)d\theta + \int_{\underline{\theta}}^{\overline{\theta}}\{I(\theta,q(\theta)) - [v(\theta,q(\theta)) - u(\theta)]\}d\Phi(\theta)$$

$$= \int_{\underline{\theta}}^{\overline{\theta}}[v(\theta,q(\theta)) - c(q(\theta))]f(\theta)d\theta + \int_{\underline{\theta}}^{\overline{\theta}}[I(\theta,q(\theta)) - v(\theta,q(\theta))]d\Phi(\theta) - u(\underline{\theta}) - \int_{\underline{\theta}}^{\overline{\theta}} v_\theta(\theta,q\,(\theta))d\theta$$

$$+\int_{\underline{\theta}}^{\overline{\theta}} F(\theta)v_\theta(\theta,q(\theta))d\theta + u(\underline{\theta})[\Phi(\overline{\theta})-\Phi(\underline{\theta})] + \Phi(\overline{\theta})\int_{\underline{\theta}}^{\overline{\theta}} v_\theta(\theta,q(\theta))d\theta - \int_{\underline{\theta}}^{\overline{\theta}}\Phi(\theta)v_\theta(\theta,q(\theta))d\theta,$$

which, by collecting terms, can be simplified to further obtain

$$\int_{\underline{\theta}}^{\overline{\theta}}[v(\theta,q(\theta)) - c(q(\theta))]f(\theta)d\theta + \int_{\underline{\theta}}^{\overline{\theta}}\left[\frac{F(\theta)-\Phi(\theta)+\Phi(\overline{\theta})-1}{f(\theta)}\right]v_\theta(\theta,q(\theta))f(\theta)d\theta$$

$$+\int_{\underline{\theta}}^{\overline{\theta}}\frac{\phi(\theta)[I(\theta,q(\theta)) - v(\theta,q(\theta))]}{f(\theta)}f(\theta)d\theta + u(\underline{\theta})[\Phi(\overline{\theta})-\Phi(\underline{\theta})-1].$$

By collecting terms one more time and dropping irrelevant constants, this expression reduces to that in (6). The following result is the analogue of Result 4 in the Supplementary Appendix.

Result 1. *Under (PS), (BCH), and (FP), the implementable allocation $\{u(\theta), q(\theta)\}$ solves the simple BC problem if, and only if, there exists a cumulative multiplier function $\Phi(\theta)$ on $[\underline{\theta}, \overline{\theta}]$ such that the first-order conditions (7) and the complementary slackness condition (8) are satisfied. Moreover, $q(\theta)$ is continuous.*

We now turn to proving Proposition 1. Consider a solution to the IR problem. We claim that under sufficient conditions, it is also a solution to the BC problem. For notational simplicity, in the following we suppress the subscript *IR* from $u_{IR}(\theta)$, $q_{IR}(\theta)$, $t_{IR}(\theta)$, and $\overline{u}_{IR}(\theta)$. To start, by Result 4 in the Supplementary Appendix, an implementable allocation $\{u(\theta), q(\theta)\}$ solves the IR problem if, and only if, there exists a cumulative multiplier function $\gamma(\theta)$ with the properties of a cumulative distribution function such that the first-order conditions

$$v_q(\theta, q(\theta)) - c'(q(\theta)) = \frac{\gamma(\theta) - F(\theta)}{f(\theta)} v_{\theta q}(\theta, q(\theta)) \tag{26}$$

for each type and the complementary slackness condition

$$\int_{\underline{\theta}}^{\overline{\theta}} [u(\theta) - \overline{u}(\theta)]\, d\gamma(\theta) = 0 \tag{27}$$

hold, together with $u(\theta) \geq \overline{u}(\theta)$. By Result 1 above, the allocation that solves the IR problem solves the BC problem if, and only if, there exists a cumulative multiplier function $\Phi(\theta)$ such that the first-order conditions

$$v_q(\theta, q(\theta)) - c'(q(\theta)) = \left[\frac{\Phi(\theta) - F(\theta) + 1 - \Phi(\overline{\theta})}{f(\theta)}\right] v_{\theta q}(\theta, q(\theta)) + \frac{\phi(\theta)\,[v_q(\theta, q(\theta)) - I_q(\theta, q(\theta))]}{f(\theta)} \tag{28}$$

for each type and the complementary slackness condition

$$\int_{\underline{\theta}}^{\overline{\theta}} [I(\theta, q(\theta)) - v(\theta, q(\theta)) + u(\theta)]\, d\Phi(\theta) = 0 \tag{29}$$

hold, together with $t(\theta) \leq I(\theta, q(\theta))$ and $u(\theta) \geq \overline{u}$. Note that for $\Phi(\theta)$ to be a legitimate cumulative multiplier, it must be nonnegative and weakly increasing with θ. Let $\Phi(\theta) = \alpha + \gamma(\theta)$ be the cumulative multiplier in the BC problem. Clearly, $\Phi(\theta) = \alpha + \gamma(\theta)$ for any constant α is a legitimate cumulative multiplier. Also, with $\Phi(\theta) = \alpha + \gamma(\theta)$, the multiplier $d\gamma(\theta)$ on the IR constraint of type θ is zero or strictly positive if, and only if, the multiplier $d\Phi(\theta)$ on the BC constraint of type θ is zero or strictly positive.

The rest of the proof proceeds in three steps. In the first step, we show that at the IR allocation, the complementary slackness condition of the BC problem, (29), holds and the IR allocation satisfies $t(\theta) \leq I(\theta, q(\theta))$ and $u(\theta) \geq \overline{u}$. In the second step, we argue that the first-order conditions of the BC problem in (28) are identical to those of the IR problem in (26). In the third step, we show that consumers reach the same utility in the two problems.

Step 1: Verify Complementary Slackness Condition in BC Problem, $t(\theta) \leq I(\theta, q(\theta))$, and $\overline{u}(\theta) \geq \overline{u}$. We now claim that at the IR allocation, the complementary slackness condition in the BC problem, (29), holds and the IR allocation satisfies $t(\theta) \leq I(\theta, q(\theta))$. To this purpose, recall that $I(\theta, q(\theta)) \geq v(\theta, q(\theta)) - \overline{u}(\theta)$, $I(\theta, q(\theta)) = v(\theta, q(\theta)) - \overline{u}(\theta)$, and $\overline{u}(\underline{\theta}) \geq \overline{u}$ by assumption. Note that when the IR constraints bind so that $d\gamma(\theta) = d\Phi(\theta) > 0$, then $q(\theta) = \overline{q}(\theta)$ and $v(\theta, q(\theta)) - t(\theta) = \overline{u}(\theta)$ or, equivalently, $t(\theta) = v(\theta, q(\theta)) - \overline{u}(\theta)$. Since, by assumption, $v(\theta, q(\theta)) - \overline{u}(\theta) = I(\theta, q(\theta))$ for types whose IR constraints bind, it follows that $t(\theta) = I(\theta, q(\theta))$. When, instead, the IR constraints do not bind so that $d\gamma(\theta) = d\Phi(\theta) = 0$, then $v(\theta, q(\theta)) - t(\theta) \geq \overline{u}(\theta)$ or, equivalently, $t(\theta) \leq v(\theta, q(\theta)) - \overline{u}(\theta)$. Since, by assumption, $v(\theta, q(\theta)) - \overline{u}(\theta) \leq I(\theta, q(\theta))$ for consumers whose IR constraints do not bind, it follows that $t(\theta) \leq I(\theta, q(\theta))$. Hence, if condition (27) holds for the IR problem, then condition (29) holds for the BC problem. Also, $t(\theta) \leq I(\theta, q(\theta))$ is satisfied, and $u(\theta) \geq \overline{u}$ with strict inequality for any $\theta > \underline{\theta}$, since $u(\theta) \geq \overline{u}(\theta)$, $\overline{u}(\theta)$ is strictly increasing, and, by assumption, $\overline{u}(\underline{\theta}) \geq \overline{u}$.

Step 2: Verify First-Order Conditions of IR Problem Identical to Those of BC Problem. We now show that given the cumulative multiplier $\Phi(\theta)$, the quantity profile that solves the IR problem satisfies the first-order conditions of the BC problem, namely (28). First, note that

$$\Phi(\theta) + 1 - \Phi(\overline{\theta}) = \alpha + \gamma(\theta) + 1 - [\alpha + \gamma(\overline{\theta})] = \gamma(\theta), \tag{30}$$

where the first equality follows from $\Phi(\theta) = \alpha + \gamma(\theta)$ and the second equality holds since $\gamma(\overline{\theta}) = 1$; see the Supplementary Appendix for a proof of this latter result. Second, observe that, by assumption, $I_q(\theta, q(\theta))$ equals $v_q(\theta, q(\theta))$ when the IR constraints bind. Thus, for each θ, either $\phi(\theta) = 0$ or, if not, $I_q(\theta, q(\theta)) = v_q(\theta, q(\theta))$. Hence, the second term on the right side of (28) equals zero for each θ. These two observations imply that the first-order conditions of the BC problem in (28) are identical to those of the IR problem in (26).

Step 3: Same Utility in IR and BC Problems. The requirement that $I(\theta, \overline{q}(\theta)) = v(\theta, \overline{q}(\theta)) - \overline{u}(\theta)$ for types whose

IR constraints bind in the IR problem ensures that the utility achieved by each consumer is identical in the IR and BC problems. Specifically, let θ' be such type. Then, for any type θ higher than θ', in the IR problem we have

$$u(\theta) = \overline{u}(\theta') + \int_{\theta'}^{\theta} v_\theta(x, q(x))dx = v(\theta', \overline{q}(\theta')) - I(\theta', \overline{q}(\theta')) + \int_{\theta'}^{\theta} v_\theta(x, q(x))dx, \tag{31}$$

since $u'(\theta) = v_\theta(x, q(x))$ by local incentive compatibility, and $u(\theta') = \overline{u}(\theta') = v(\theta', \overline{q}(\theta')) - I(\theta', \overline{q}(\theta'))$ by assumption. The utility in (31) equals the utility that the consumer achieves in the solution to the BC problem, given that the BC constraints bind in the BC problem if, and only if, the IR constraints bind in the IR problem and the optimal quantity profiles in the two problems coincide, as argued in Step 2. An analogous argument holds for any type lower than θ'. Hence, consumers' utility schedules coincide in the two problems. Lastly, $u(\theta) \geq \overline{u}$ since $\overline{u}(\underline{\theta}) \geq \overline{u}$ by assumption so that full participation is satisfied in both problems.

Thus, the solutions to the IR and BC problems are the same. By an argument similar to the one in the proof of Result 1 in the Supplementary Appendix, it is also possible to show that $\Phi(\overline{\theta}) = 1$. $\square$

Proof of Proposition 2: Recall that $I(\theta, q) = Y - \underline{z}(\theta, q)$, where

$$\underline{z}(\theta, q) = -\underline{z}_1(\theta) - z_2\nu(q),\ \underline{z}_1'(\theta) = \psi(\log(\theta - z_2)), \text{ and } \underline{\theta} > z_2 > 0. \tag{32}$$

Let $\psi(\cdot)$ be a positive continuous function. To show that (BCH) is satisfied under these assumptions, we proceed by showing that it is possible to construct a menu $\{\bar{t}(\theta), \overline{q}(\theta)\}$ such that $\bar{t}(\theta) = I(\theta, \overline{q}(\theta)) = Y - \underline{z}(\theta, \overline{q}(\theta))$, $\bar{t}'(\theta) = \theta\nu'(\overline{q}(\theta))\overline{q}'(\theta)$, and $\overline{q}(\theta)$ is weakly increasing. Since this third condition can be restated as $\overline{\theta}(q)$ is weakly increasing, it is possible to define the function $\overline{T}(q)$ such that $\overline{T}(q) = \bar{t}(\overline{\theta}(q))$. Thus, the first requirement of (BCH) amounts to $\overline{T}(q) = I(\overline{\theta}(q), q) = Y - \underline{z}(\overline{\theta}(q), q)$, whereas the second requirement amounts to $\bar{t}'(\overline{\theta}(q))\overline{\theta}'(q) = \overline{\theta}(q)\nu'(q)$ or, equivalently, $\overline{T}'(q) = \overline{\theta}(q)\nu'(q)$, with $q = \overline{q}(\theta) \in [\overline{q}(\underline{\theta}), \overline{q}(\overline{\theta})]$. Then, rather than establishing that we can construct an increasing function $\overline{q}(\theta)$ that satisfies $\bar{t}(\theta) = Y - \underline{z}(\theta, \overline{q}(\theta))$ and $\bar{t}'(\theta) = \theta\nu'(\overline{q}(\theta))\overline{q}'(\theta)$ under (32), we show, equivalently, that we can construct an increasing function $\overline{\theta}(q)$ that satisfies

$$\begin{cases} \overline{T}(q) = Y - \underline{z}(\overline{\theta}(q), q) \\ \overline{T}'(q) = \overline{\theta}(q)\nu'(q) \end{cases} \tag{33}$$

under (32). Now, using (32) it follows that the derivative of the first expression in (33) with respect to q is

$$\overline{T}'(q) = \underline{z}_1'(\overline{\theta}(q))\overline{\theta}'(q) + z_2\nu'(q) = \psi(\log[\overline{\theta}(q) - z_2])\overline{\theta}'(q) + z_2\nu'(q).$$

By equating the right sides of this last expression and of the second expression in (33), we obtain

$$\psi(\log[\overline{\theta}(q) - z_2])\overline{\theta}'(q) + z_2\nu'(q) = \overline{\theta}(q)\nu'(q) \Leftrightarrow \nu'(q) = \psi(\log[\overline{\theta}(q) - z_2])\overline{\theta}'(q)/[\overline{\theta}(q) - z_2]. \tag{34}$$

By integrating both sides of (34) from $\overline{q}(\underline{\theta})$ to $q \leq \overline{q}(\overline{\theta})$ and using $\underline{\theta} = \overline{\theta}(\overline{q}(\underline{\theta}))$, it follows that

$$\nu(q) - \nu(\overline{q}(\underline{\theta})) = \Psi(\log[\overline{\theta}(q) - z_2]) - \Psi(\log(\underline{\theta} - z_2)), \tag{35}$$

where $\Psi(\cdot)$, the integral of $\psi(\cdot)$, is weakly increasing since $\psi(\cdot)$ is positive. Simple manipulations yield

$$\overline{\theta}(q) = z_2 + \exp\{(\Psi)^{-1}(\nu(q) - \nu(\overline{q}(\underline{\theta})) + \Psi(\log(\underline{\theta} - z_2)))\},$$

with $\overline{q}(\underline{\theta})$ determined by the last equality in (34) evaluated at $q = \overline{q}(\underline{\theta})$. Note that $\overline{\theta}(q)$ is an increasing function of q, since $(\Psi)^{-1}(\cdot)$ and $\nu(\cdot)$ are increasing functions. So, $\overline{q}(\theta)$ is an increasing function of θ. Moreover, $\overline{T}(q) = Y + \underline{z}_1(\overline{\theta}(q)) + z_2\nu(q)$, which implies

$$\overline{T}'(q) = \underline{z}_1'(\overline{\theta}(q))\overline{\theta}'(q) + z_2\nu'(q) = \nu'(q)[\overline{\theta}(q) - z_2] + z_2\nu'(q) = \overline{\theta}(q)\nu'(q),$$

where the second equality follows from (34). So the three requirements of (BCH) are satisfied, and indeed $T'(q) = \theta(q)\nu'(q)$ for types whose BC constraints bind. For example, it is easy to show that if $\underline{z}(\theta, q) = z_0 - z_1(\theta - z_2)^\lambda -$

$z_2\nu(q)$, $\lambda \geq 2$, $z_1, z_2 > 0$, and $\underline{\theta} > z_2$, then

$$\overline{q}(\theta) = (\nu)^{-1}\left\{\nu(\overline{q}(\underline{\theta})) + \frac{\lambda z_1}{\lambda-1}\left[(\theta - z_2)^{\lambda-1} - (\underline{\theta} - z_2)^{\lambda-1}\right]\right\},\ \overline{q}'(\theta) > 0,\ \text{and}$$

$$\overline{T}(q) = Y(w) - z_0 + z_1\left\{(\underline{\theta} - z_2)^{\lambda-1} + \frac{\lambda-1}{\lambda z_1}\left[\nu(q) - \nu(\overline{q}(\underline{\theta}))\right]\right\}^{\frac{\lambda}{\lambda-1}} + z_2\nu(q). \quad \square$$

The Two-Dimensional Case: Suppose that the parameter w differs across consumers so that the budget schedule is $I(\theta, q, w) = Y(w) - \underline{z}(\theta, q)$. The analysis of this case differs from that of the case of constant w depending on whether the seller can discriminate across consumers based on w (Case 1) or, rather, only based on a menu of prices at most contingent on q (Case 2).

Case 1: Contractible Income Characteristic. Suppose that the seller can segment consumers across submarkets indexed by w and offer nonlinear prices in each submarket w so as to screen consumers based on θ. For ease of exposition, suppose that there are only two levels of w, say, w_L and w_H, with $Y(w_H) > Y(w_L)$. In any such submarket w, the seller's problem is as stated in the BC problem with income $Y(w)$ and budget schedule $I(\theta, q, w)$. For the corresponding simple BC problem, the necessary and sufficient conditions for an optimal solution are given by the analogue of Result 1 under the same maintained assumptions: the implementable allocation $\{u(\theta, w), q(\theta, w)\}$ solves the simple BC problem if, and only if, there exists a cumulative multiplier function $\Phi(\theta, w)$ such that the first-order conditions in (28) and the complementary slackness condition in (29) apply with $I(\theta, q, w) = Y(w) - \underline{z}(\theta, q)$. Our next result shows how this menu varies across submarkets. For this, let

$$t(\theta, w_H) = t(\theta, w_L) + Y(w_H) - Y(w_L),\ q(\theta, w_H) = q(\theta, w_L),\ \text{and}\ \Phi(\theta, w_H) = \Phi(\theta, w_L). \tag{36}$$

Result 2. *If $\{u(\theta, w_L), q(\theta, w_L)\}$ with associated cumulative multipliers $\{\Phi(\theta, w_L)\}$ solves the simple BC problem in submarket w_L, then $\{u(\theta, w_H), q(\theta, w_H)\}$ with associated cumulative multipliers $\{\Phi(\theta, w_H)\}$ satisfying (36) solves the simple BC problem in submarket w_H.*

This result states that type (θ, w_H) in the submarket with the higher income level is offered the same quantity as type (θ, w_L) in the submarket with the lower income level, that is, $q(\theta, w_H) = q(\theta, w_L)$. Moreover, the binding patterns of the multipliers in the two submarkets are identical in that the cumulative multiplier binds for type (θ, w_H) in submarket w_H if, and only if, it binds for type (θ, w_L) in submarket w_L. The only difference is that type (θ, w_H) in submarket w_H pays $Y(w_H) - Y(w_L)$ more for the same quantity purchased by type (θ, w_L) in submarket w_L. The idea is straightforward. In the submarket with income $Y(w_L)$, a consumer with taste θ chooses the pair $(t(\theta, w_L), q(\theta, w_L))$ leading to the consumption of $z(\theta, w_L) = Y(w_L) - t(\theta, w_L)$ units of the numeraire good. The consumption bundle $(q(\theta, w_L), z(\theta, w_L))$ must jointly provide enough calories so that the consumer meets the constraint $z(\theta, w_L) \geq \underline{z}(\theta, q(\theta, w_L))$. Suppose that this constraint binds for a consumer with taste θ, that is,

$$z(\theta, w_L) = \underline{z}(\theta, q(\theta, w_L)) = Y(w_L) - t(\theta, w_L). \tag{37}$$

In submarket w_H, at $(t(\theta, w_L), q(\theta, w_L))$ the budget constraint is slack for a consumer with taste θ since $Y(w_H) > Y(w_L)$. Clearly, in submarket w_H, it is feasible for the seller to offer the same quantity as in submarket w_L, that is, $q(\theta, w_H) = q(\theta, w_L)$, since $q(\theta, w_L)$ is implementable in submarket w_H too, and simply increase the price by $Y(w_H) - Y(w_L)$. In the proof of Result 2, we show that doing so is in general optimal for the seller.

Proof of Result 2: Let $\{u(\theta, w_L), q(\theta, w_L)\}$ and the cumulative multipliers $\{\Phi(\theta, w_L)\}$ solve the simple BC problem in submarket w_L. By Result 1, we know that these schedules satisfy the first-order conditions (28) and the complementary slackness condition (29) with $t(\theta)$, $q(\theta)$, $\Phi(\theta)$, $\phi(\theta)$, and $I(\theta, q)$ replaced by $t(\theta, w_L)$, $q(\theta, w_L)$, $\Phi(\theta, w_L)$, $\phi(\theta, w_L)$, and $I(\theta, q, w_L)$. It is immediate that the allocations and multipliers given in (36) solve the corresponding first-order and complementary slackness conditions for submarket w_H. To see why, note that since $I_q(\theta, q, w) = -\underline{z}_q(\theta, q)$ is independent of w (conditional on q), the first-order conditions in the two submarkets are identical under (36). Consider next the complementary slackness condition. Since this condition holds in submarket w_L, for any θ whose budget constraint for the seller's good binds and so $\phi(\theta, w_L)$ is positive, we have

$$t(\theta, w_L) = I(\theta, q(\theta, w_L), w_L) \equiv Y(w_L) - \underline{z}(\theta, q(\theta, w_L)). \tag{38}$$

But then for this same θ in submarket w_H under (36), the multiplier $\phi(\theta, w_H)$ is also positive, since

$$t(\theta, w_H) = t(\theta, w_L) + Y(w_H) - Y(w_L) = Y(w_H) - \underline{z}(\theta, q(\theta, w_L)) = Y(w_H) - \underline{z}(\theta, q(\theta, w_H)),$$

where the first and third equalities follow from (36), and the second equality from (38). Hence, the conjectured solution satisfies the first-order conditions and complementary slackness condition for submarket w_H. So, by Result 1, this conjectured solution solves the simple BC problem for submarket w_H. □

Case 2: Noncontractible Income Characteristic. Suppose now that the seller cannot segment consumers across submarkets. That is, the seller must offer the same price schedule to all consumers regardless of their w (and θ). This environment is equivalent to one in which the seller observes neither w nor θ. Assume that w and θ are sufficiently positively dependent that w can be expressed as a nonlinear monotone function of θ, namely, $w = \omega(\theta)$ with $\omega'(\theta) > 0$. Then, substituting $w = \omega(\theta)$ into $I(\theta, q, w) = Y(w) - \underline{z}(\theta, q)$ gives

$$I(\theta, q, \omega(\theta)) = Y(\omega(\theta)) - \underline{z}(\theta, q). \tag{39}$$

Under (39), the analogues of Proposition 1 and Result 1 apply. To see that the analogue of Proposition 2 also holds, let $Y(\omega(\theta)) = Y + y(\omega(\theta))$ without loss. Then, the analogous result holds with $v(\theta, q) = \theta\nu(q)$, $\underline{z}(\theta, q) = -\underline{z}_1(\theta) - z_2\nu(q)$, and $y'(\omega(\theta))\omega'(\theta) + \underline{z}_1'(\theta) = \psi(\log(\theta - z_2))$ for $\underline{\theta} > z_2 > 0$. □

Proof of Proposition 3: Recall that $T'(q(\theta)) = \theta\nu'(q(\theta)) > 0$ by local incentive compatibility, and $A(q) = -\nu''(q)/\nu'(q)$. Differentiating $T'(q) = \theta(q)\nu'(q)$ yields

$$T''(q) = \theta'(q)\nu'(q) + \theta(q)\nu''(q) = \theta(q)\nu'(q)\left[\frac{\theta'(q)}{\theta(q)} + \frac{\nu''(q)}{\nu'(q)}\right] = T'(q)\left[\frac{1}{\theta(q)q'(\theta)} - A(q)\right]. \tag{40}$$

By using $T'(q) = \theta(q)\nu'(q)$, the first-order condition (9) can be expressed as $\{\theta - [\gamma(\theta) - F(\theta)]/f(\theta)\}\nu'(q(\theta)) - c = 0$. Applying the implicit function theorem to this condition, we obtain

$$q'(\theta) = -\frac{\frac{\partial}{\partial\theta}\left[\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}\right]\nu'(q(\theta))}{\left[\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}\right]\nu''(q(\theta))} = \frac{\frac{\partial}{\partial\theta}\left[\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}\right]}{\left[\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}\right]A(q(\theta))}.$$

Note that $\theta > [\gamma(\theta) - F(\theta)]/f(\theta)$, since the seller's first-order condition can also be expressed as $\{\theta - [\gamma(\theta) - F(\theta)]/f(\theta)\} = c/\nu'(q(\theta))$ and both c and $\nu'(q(\theta))$ are strictly positive by assumption. Also, $q'(\theta) > 0$ by assumption (PS); see Jullien (2000). Using (40) and the fact that $T'(q), A(q) > 0$, we can equivalently express $T''(q) \leq 0$ as

$$T'(q(\theta))A(q(\theta))\left\{\frac{\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}}{\theta\frac{\partial}{\partial\theta}\left[\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}\right]} - 1\right\} \leq 0 \Leftrightarrow \frac{\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}}{\theta\frac{\partial}{\partial\theta}\left[\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}\right]} \leq 1. \tag{41}$$

We establish the desired result by showing that (41) holds.

Consider first any interval of consumer types whose IR constraints bind. By construction, any such type θ consumes $\overline{q}(\theta)$ and achieves utility $\overline{u}(\theta)$. The homogeneity assumption implies $\nu(\overline{q}(\theta)) - \overline{u}'(\theta) = 0$, which in turn yields

$$\overline{q}'(\theta) = -\frac{-\overline{u}''(\theta)}{\nu'(\overline{q}(\theta))} = \frac{\overline{u}''(\theta)}{\nu'(\overline{q}(\theta))}$$

by applying the implicit function theorem, with $\overline{q}'(\theta) > 0$ by assumption, and so by (40)

$$T''(\overline{q}) = \frac{\nu'(\overline{q})}{\overline{q}'(\theta)} + \theta(\overline{q})\nu''(\overline{q}) = \frac{[\nu'(\overline{q})]^2}{\overline{u}''(\theta)} + \theta(\overline{q})\nu''(\overline{q}) = \nu'(\overline{q})\left\{\frac{\nu'(\overline{q})}{\overline{u}''(\theta)} - \theta(\overline{q})A(\overline{q})\right\}.$$

Since $\nu'(\overline{q})>0$ by assumption, it follows $T''(\overline{q})\leq 0$ if, and only if, $\nu'(\overline{q})\leq\theta(\overline{q})A(\overline{q})\overline{u}''(\theta)$, which holds by assumption.

Consider now any interval of consumer types whose IR constraints do not bind, in which case $\gamma(\theta) = \gamma$ for all types in such an interval. When $\gamma = 1$, it follows that

$$q'(\theta) = \frac{\frac{\partial}{\partial\theta}\left[\theta - \frac{1-F(\theta)}{f(\theta)}\right]}{\left[\theta - \frac{1-F(\theta)}{f(\theta)}\right] A(q(\theta))} \geq \frac{1}{\theta A(q(\theta))} = \frac{\nu'(q(\theta))}{-\theta\nu''(q(\theta))}, \tag{42}$$

where the inequality in the above follows from the assumption that $[1 - F(\theta)]/f(\theta)$ decreases with θ. Condition (42) implies $1/q'(\theta) \leq -\theta(q)\nu''(q)/\nu'(q)$, which combined with (40) yields

$$T''(q) = \frac{\nu'(q)}{q'(\theta)} + \theta(q)\nu''(q) \leq \nu'(q)\left[\frac{-\theta(q)\nu''(q)}{\nu'(q)}\right] + \theta(q)\nu''(q) = 0.$$

When, instead, $\gamma \in [0, 1)$, the last inequality in (41) becomes

$$\frac{\theta f(\theta) - \gamma + F(\theta)}{\theta f(\theta)} \leq \frac{\partial}{\partial\theta}\left[\theta - \frac{\gamma}{f(\theta)} + \frac{F(\theta)}{f(\theta)}\right] \Leftrightarrow \theta f^2(\theta) \geq -[\gamma - F(\theta)]f(\theta) - [\gamma - F(\theta)]\theta f'(\theta). \tag{43}$$

We prove that (43) holds by considering two further cases.

Case 1: $\gamma \geq F(\theta)$. In this case, $[\gamma - F(\theta)]f(\theta) \geq 0$ so that a sufficient condition for (43) is

$$f^2(\theta) \geq -[\gamma - F(\theta)]f'(\theta). \tag{44}$$

If $f'(\theta) \geq 0$, then it is immediate that the displayed inequality is satisfied. Suppose now that $f'(\theta) < 0$. Since

$$\frac{\partial}{\partial\theta}\left(\frac{1 - F(\theta)}{f(\theta)}\right) = \frac{-f^2(\theta) - [1 - F(\theta)]\, f'(\theta)}{f^2(\theta)} \leq 0 \Leftrightarrow f^2(\theta) \geq -\left[1 - F(\theta)\right] f'(\theta)$$

by assumption and $-[1 - F(\theta)]f'(\theta) > -[\gamma - F(\theta)]f'(\theta)$ with $\gamma < 1$ and $f'(\theta) < 0$, it follows that (44) and so (43) are satisfied.

Case 2: $\gamma < F(\theta)$. In this case, we can rewrite (43) as

$$\theta f^2(\theta) \geq [F(\theta) - \gamma]f(\theta) + [F(\theta) - \gamma]\theta f'(\theta) \Leftrightarrow \theta f(\theta) \geq F(\theta) - \gamma + [F(\theta) - \gamma]\theta\frac{f'(\theta)}{f(\theta)}. \tag{45}$$

When $f'(\theta) \leq 0$, a sufficient condition for the last displayed inequality to hold is $f(\theta) \geq F(\theta) - \gamma$, which is satisfied whenever $\underline{\theta} f(\underline{\theta}) \geq 1$ since $F(\theta) - \gamma \leq 1$. But since $\underline{\theta} f(\underline{\theta}) \geq 1$ holds by assumption, it follows that (43) is satisfied. When, instead, $f'(\theta) > 0$, a sufficient condition for (45) is

$$\theta f(\theta) \geq F(\theta) + F(\theta)\theta\frac{f'(\theta)}{f(\theta)} \Leftrightarrow f^2(\theta) \geq \frac{F(\theta)f(\theta)}{\theta} + F(\theta)f'(\theta)$$

or, equivalently,

$$f^2(\theta) - F(\theta)f'(\theta) \geq \frac{f(\theta)}{f(\theta)}\frac{F(\theta)f(\theta)}{\theta} = \frac{f^2(\theta)F(\theta)}{\theta f(\theta)} \Leftrightarrow \frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right) = \frac{f^2(\theta) - F(\theta)f'(\theta)}{f^2(\theta)} \geq \frac{F(\theta)}{\theta f(\theta)},$$

which can be restated as $\varepsilon_r(\theta) = \frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right)\theta\left[\frac{F(\theta)}{f(\theta)}\right]^{-1} \geq 1$, which is satisfied by assumption. □

Proof of Proposition 4. Let the allocation in the standard nonlinear pricing model be $\{u_s(\theta), q_s(\theta)\}$. A seller's first-order conditions in the standard model and the augmented model can be written, respectively, as

$$1 - \frac{c}{T'(q_s(\theta))} = \frac{1 - F(\theta)}{\theta f(\theta)} \text{ and } 1 - \frac{c}{T'(q(\theta))} = \frac{\gamma(\theta) - F(\theta)}{\theta f(\theta)}. \tag{46}$$

Recall that in the standard model, the IR constraints bind only for the lowest type so that $u_s(\underline{\theta}) = \overline{u}$. Since $\gamma(\theta) \leq 1$, it follows that

$$[1 - F(\theta)]/\theta f(\theta) \geq [\gamma(\theta) - F(\theta)]/\theta f(\theta)$$

with strict inequality if $\gamma(\theta) < 1$. Using (46) and local incentive compatibility then yields

$$T'(q(\theta)) = \theta\nu'(q(\theta)) \leq T'(q_s(\theta)) = \theta\nu'(q_s(\theta)). \tag{47}$$

Thus, $T'(q(\theta)) \leq T'(q_s(\theta))$ with strict inequality if $\gamma(\theta) < 1$, which further implies that $q(\theta) \geq q_s(\theta)$ since $\nu'(\cdot)$ is decreasing by assumption. This latter result, in turn, yields that

$$u(\theta) = u(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u'(x)dx = u(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \nu(q(x))dx \geq u_s(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \nu(q_s(x))dx = u_s(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u_s'(x)dx = u_s(\theta), \tag{48}$$

where the second and third equalities in (48) follow from local incentive compatibility in both the standard and the augmented model in that $u'(\theta) = \nu(q(\theta))$ and $u_s'(\theta) = \nu(q_s(\theta))$. The inequality in (48) follows because $u(\underline{\theta}) \geq \overline{u}(\underline{\theta})$, $\overline{u}(\underline{\theta}) \geq \overline{u}$ by assumption, $u_s(\underline{\theta}) = \overline{u}$ since the IR constraint binds at $\underline{\theta}$ in the standard model, $q(\theta) \geq q_s(\theta)$ as just argued, and $\nu(\cdot)$ is strictly increasing. □

Proof of Proposition 5: We divide the proofs into two parts, Case a) and Case b). In both parts, we rely on the assumption of full participation under nonlinear and linear pricing.

Case a). We start by showing that if the price schedule exhibits quantity discounts in that $p'(q) \leq 0$ at $q = q(\theta)$ and if $q_m(\theta) \geq q(\theta)$, then the utility of a consumer of type θ is higher under linear than under nonlinear pricing, that is, $u_m(\theta) \geq u(\theta)$. By contradiction, assume that $p'(q) \leq 0$ at $q = q(\theta)$ and $q_m(\theta) \geq q(\theta)$ but

$$u(\theta) = \theta\nu(q(\theta)) - T(q(\theta)) > u_m(\theta) = \theta\nu(q_m(\theta)) - \theta\nu'(q_m(\theta))q_m(\theta), \tag{49}$$

where in (49) we have used the fact that $p_m = \theta\nu'(q_m(\theta))$ under linear pricing by consumer optimality. Given that $q_m(\theta)$ maximizes the consumer's utility under linear pricing, it follows that

$$\theta\nu(q(\theta)) - T(q(\theta)) > \theta\nu(q_m(\theta)) - \theta\nu'(q_m(\theta))q_m(\theta) \geq \theta\nu(q(\theta)) - \theta\nu'(q_m(\theta))q(\theta), \tag{50}$$

which implies $\theta\nu(q(\theta)) - T(q(\theta)) > \theta\nu(q(\theta)) - \theta\nu'(q_m(\theta))q(\theta)$ or, equivalently,

$$\theta\nu'(q_m(\theta)) > T(q(\theta))/q(\theta) = p(q(\theta)). \tag{51}$$

Note that the first inequality in (50) restates (49), whereas the second inequality follows from the fact that any quantity demanded that is different from $q_m(\theta)$, including $q(\theta)$, implies a lower utility for the consumer at the linear price p_m. Next, by the assumption of the case, $p'(q) = [T'(q) - T(q)/q]/q \leq 0$ or, equivalently,

$$T'(q(\theta)) \leq T(q(\theta))/q(\theta). \tag{52}$$

This inequality, together with (51) and the incentive compatibility condition $T'(q(\theta)) = \theta\nu'(q(\theta))$, implies

$$\theta\nu'(q(\theta)) = T'(q(\theta)) \leq T(q(\theta))/q(\theta) < \theta\nu'(q_m(\theta)), \tag{53}$$

and so $\theta\nu'(q_m(\theta)) > \theta\nu'(q(\theta))$, which is a contradiction since $q_m(\theta) \geq q(\theta)$ by assumption and $\nu'(\cdot)$ is decreasing. Hence, $u_m(\theta) \geq u(\theta)$.

Case b). We now show that if the price schedule exhibits quantity discounts in that $T''(q) \leq 0$ for all $q = q(\theta)$, $\gamma(\theta) < 1$, and $q(\theta) > q_m(\theta)$, then the utility of a consumer of type θ is higher under linear pricing than under nonlinear pricing. Consider one such type, say, $\hat{\theta}$. By way of contradiction, suppose that $u(\hat{\theta}) > u_m(\hat{\theta})$. We will show that if this assumption holds, then we contradict the assumption that all types participate under linear pricing. That is, if $u(\hat{\theta}) > u_m(\hat{\theta})$, then there exists a type $\theta_2 > \hat{\theta}$ whose participation constraint binds under nonlinear pricing, that is, $u(\theta_2) = \bar{u}(\theta_2)$, but is violated under linear pricing, that is, $u_m(\theta_2) < \overline{u}(\theta_2)$. Hence, type θ_2 is excluded under linear pricing. Note that a consumer $\theta_2 > \hat{\theta}$ with $u(\theta_2) = \bar{u}(\theta_2)$ must exist if $\gamma(\hat{\theta}) < 1$.

To reach the desired contradiction, rewrite $u(\hat{\theta}) > u_m(\hat{\theta})$ as

$$u(\theta_2) - [u(\theta_2) - u(\hat{\theta})] > u_m(\theta_2) - [u_m(\theta_2) - u_m(\hat{\theta})], \tag{54}$$

which can be equivalently expressed as

$$u(\theta_2) - \int_{\hat{\theta}}^{\theta_2} u'(x)dx > u_m(\theta_2) - \int_{\hat{\theta}}^{\theta_2} u'_m(x)dx. \tag{55}$$

Condition (55), in turn, is equivalent to

$$\overline{u}(\theta_2) - u_m(\theta_2) > \int_{\hat{\theta}}^{\theta_2} \left[\nu(q(x)) - \nu(q_m(x))\right] dx \tag{56}$$

by using $u(\theta_2) = \overline{u}(\theta_2)$ since the IR constraint binds at θ_2 under nonlinear pricing by construction, by exploiting incentive compatibility under nonlinear and linear pricing, namely, $u'(\theta) = \nu(q(\theta))$ and $u'_m(\theta) = \nu(q_m(\theta))$ for all types, and by rearranging terms.

We now argue that the right side of (56) is positive, which establishes the desired contradiction. To see that the right side of (56) is positive, note first that for all $\theta \geq \hat{\theta}$,

$$p_m = \theta\nu'(q_m(\theta)) = \hat{\theta}\nu'(q_m(\hat{\theta})) \geq \hat{\theta}\nu'(q(\hat{\theta})) = T'(q(\hat{\theta})) \geq T'(q(\theta)) = \theta\nu'(q(\theta)), \tag{57}$$

where the first two equalities follow from a consumer's first-order condition under linear pricing, which, of course, holds for each θ, the first inequality follows from $q(\hat{\theta}) > q_m(\hat{\theta})$ by the assumption of the case and the fact that $\nu'(\cdot)$ is decreasing, the third and fourth equalities follow from local incentive compatibility under nonlinear pricing, and the second inequality holds for any $\theta \geq \hat{\theta}$ since $T''(q) \leq 0$ at all $q = q(\theta)$ by assumption. Hence, (57) implies $\theta\nu'(q_m(\theta)) \geq \theta\nu'(q(\theta))$ for all $\theta \geq \hat{\theta}$, and so $q(\theta) \geq q_m(\theta)$ for all $\theta \geq \hat{\theta}$, given that $\nu'(\cdot)$ is decreasing. But since $\nu(\cdot)$ is increasing, then the right side of (56) is positive when $q(\theta) \geq q_m(\theta)$ for all $\theta \geq \hat{\theta}$, which implies $\overline{u}(\theta_2) > u_m(\theta_2)$. Then, θ_2 does not participate under linear pricing, which is a contradiction. □

Proof of Proposition 6: To start, note that $s(\theta, \overline{q}(\theta)) \geq \overline{u}(\theta)$ for consumers with types $\theta \in [\theta', \theta'']$ implies that the seller makes nonnegative profits from each such consumer type under nonlinear pricing and these types participate by the argument in the proof of Proposition 1. To establish the desired claim, we need to show that there exists a subinterval of types in $[\theta', \theta'']$ who do not participate under linear pricing. To this purpose, suppose, by way of contradiction, that all consumer types in $[\theta', \theta'']$ participate under linear pricing. We prove that if this assumption holds, then the seller makes negative profits under linear pricing. To prove this result, let $\hat{\theta}$ be a type in $[\theta', \theta'']$ with $u_m(\hat{\theta}) = \overline{u}(\hat{\theta})$ by assumption. Observe that for any type θ in $[\hat{\theta}, \theta'']$ who participates under linear pricing, it must be $u_m(\theta) \geq \overline{u}(\theta)$, which can be expanded as

$$\begin{aligned} u_m(\theta) &= u_m(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} u'_m(x)dx = u_m(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} \nu(q_m(x))dx \\ &\geq \overline{u}(\theta) = \overline{u}(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} \overline{u}'(x)dx = \overline{u}(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} \nu(\overline{q}(x))dx, \end{aligned} \tag{58}$$

where the second equality in (58) uses the fact that $u'_m(\theta) = \nu(q_m(\theta))$ by the consumer's first-order condition under linear pricing, $\theta\nu'(q_m(\theta)) = p_m$, and the last equality uses the homogeneity assumption that $\overline{u}'(\theta) = \nu(\overline{q}(\theta))$. Since $u_m(\hat{\theta}) = \overline{u}(\hat{\theta})$ by assumption, (58) implies

$$\int_{\hat{\theta}}^{\theta} \nu(q_m(x))dx \geq \int_{\hat{\theta}}^{\theta} \nu(\overline{q}(x))dx. \tag{59}$$

With $\nu(\cdot)$ positive and increasing, (59) implies that there exists a subinterval of $[\hat{\theta}, \theta]$ with positive measure, say, $[\theta_3, \theta_4]$, such that $q_m(\theta) \geq \overline{q}(\theta)$ for all $\theta \in [\theta_3, \theta_4]$. Since $\overline{q}(\theta) > q_{FB}(\theta)$ by assumption for consumers with types in $[\theta', \theta'']$ and $q_m(\theta) \geq \overline{q}(\theta)$ for all $\theta \in [\theta_3, \theta_4]$, it follows that $q_m(\theta) > q_{FB}(\theta)$ for consumers with types in $[\theta_3, \theta_4]$. Combining

$q_m(\theta) > q_{FB}(\theta)$ with the fact that $\nu'(\cdot)$ is strictly decreasing gives

$$p_m = \theta\nu'(q_m(\theta)) < \theta\nu'(q_{FB}(\theta)) = c, \tag{60}$$

for $\theta \in [\theta_3, \theta_4]$, where the first equality follows from the first-order condition for $q_m(\theta)$ and the second equality follows from the definition of the first-best quantity for type θ, $q_{FB}(\theta)$. But $p_m < c$ implied by (60) contradicts seller optimality: the seller can always raise p_m and earn at least zero profits. □

Proof of Proposition 7: To prove this result, we exploit the equivalence between the BC and IR models. In particular, we map the change in income due to the transfer in the BC model into a corresponding change in reservation utility in the IR model. For simplicity of exposition only, we use the simple version of the equivalence between the IR and BC models by defining $\overline{u}(\theta)$ as

$$I(\theta, q) = \theta\nu(q) - \overline{u}(\theta). \tag{61}$$

Consider a consumer of type θ receiving a transfer $\tau(\theta)$ with $\tau'(\theta) \leq 0$. Hence, the consumer's budget for the seller's good is $I(\theta, q) + \tau(\theta)$ after the transfer, and the analogue of condition (61) for the equivalence between the two models under the new budget schedule is

$$I(\theta, q) + \tau(\theta) = \theta\nu(q) - \overline{u}(\theta, \tau), \tag{62}$$

where $\overline{u}(\theta, \tau)$ is the new reservation utility associated with $\tau(\theta)$. Subtracting (62) from (61) gives

$$\overline{u}(\theta, \tau) = \overline{u}(\theta) - \tau(\theta). \tag{63}$$

Hence, in the proof we need only show that replacing the original IR constraint $u(\theta) \geq \overline{u}(\theta)$, where $\overline{u}(\theta)$ is defined in (61), with the new constraint

$$u(\theta) \geq \overline{u}(\theta, \tau), \tag{64}$$

where $\overline{u}(\theta, \tau)$ is defined in (63), leads to an increase in the amount purchased of the seller's good, a decrease in marginal prices, and an increase in the total price for each quantity.

To develop some intuition, note that a consumer of type θ spends $t(\theta)$ to purchase $q(\theta)$ and the rest of her income to purchase $z = Y - t(\theta) \geq \underline{z}(\theta, q(\theta))$ before the transfer is introduced, so her budget constraint for the seller's good is $t(\theta) \leq Y - \underline{z}(\theta, q(\theta))$. If this consumer receives a transfer of $\tau(\theta)$, then the seller can ask for a higher price, since the consumer's ability to pay has increased, that is, her new budget constraint for the seller's good is $t(\theta) \leq Y + \tau(\theta) - \underline{z}(\theta, q(\theta))$. When we translate this consumer's new budget constraint for the seller's good back to a participation constraint using (62), we see from (63) that the transfer amounts to a *decrease* in the reservation utility by the amount of the transfer, which reflects the fact that the seller can now charge a higher price while still satisfying the consumer's participation constraint.

We now proceed to the formal argument. Let $\{t_\tau(\theta), q_\tau(\theta)\}$ denote the equilibrium menu with participation constraints (64) and $\{t(\theta), q(\theta)\}$ denote the original menu in the absence of the transfer with participation constraints $u(\theta) \geq \overline{u}(\theta)$. The proof involves several steps. The first step establishes that the new reservation quantity, $\overline{q}_\tau(\theta)$, is greater than the original one, $\overline{q}(\theta)$, type by type in that

$$\overline{q}_\tau(\theta) \geq \overline{q}(\theta) \text{ for all } \theta. \tag{65}$$

This result follows immediately from the fact that $\overline{q}_\tau(\theta)$ and $\overline{q}(\theta)$ satisfy assumption (BCH) in that $\overline{u}'(\theta, \tau) = \nu(\overline{q}_\tau(\theta))$ and $\overline{u}'(\theta) = \nu(\overline{q}(\theta))$, which, using (63), implies that

$$\nu(\overline{q}_\tau(\theta)) = \overline{u}'(\theta, \tau) = \overline{u}'(\theta) - \tau'(\theta) \geq \overline{u}'(\theta) = \nu(\overline{q}(\theta)), \tag{66}$$

since $\tau'(\theta) \leq 0$. As $\nu(\cdot)$ increases with q, (66) implies $\overline{q}_\tau(\theta) \geq \overline{q}(\theta)$, as desired. We show next that for each type

$$q_\tau(\theta) \geq q(\theta) \text{ and } T'_\tau(q_\tau(\theta)) \leq T'(q(\theta)). \tag{67}$$

In Part I, we establish (66) when the set of types whose IR constraints bind is the same before and after the transfer is introduced. In Part II, we establish (66) in the general case.

Part I: Suppose that the set of types whose IR constraints bind is the same before and after the transfer is introduced. Since the multiplier can bind at isolated points only at $\underline{\theta}$ or $\overline{\theta}$, we can partition $[\underline{\theta}, \overline{\theta}]$ into intervals in which the IR constraints bind for each type in the interval (Case 1) and into intervals in which they bind at either or both of the extremes of any such interval but nowhere in the interior of it (Case 2). Given this partition, intervals of Case 2 are either interior, in which case the IR constraints bind at both extremes, or they have $\underline{\theta}$ or $\overline{\theta}$ as extrema, in which case the IR constraints may or may not bind at $\underline{\theta}$ and $\overline{\theta}$.

Case 1: Consider first an interval in which the IR constraints bind for each type. Then, $\overline{q}_\tau(\theta) \geq \overline{q}(\theta)$ for all θ's in this interval by (65), and so $q_\tau(\theta) = \overline{q}_\tau(\theta) \geq \overline{q}(\theta) = q(\theta)$ for all θ and, using local incentive compatibility, $T'_\tau(q_\tau(\theta)) = \theta\nu'(q_\tau(\theta)) \leq \theta\nu'(q(\theta)) = T'(q(\theta))$ since $\nu'(\cdot)$ is decreasing, as desired.

Case 2: Consider now an interval in which the IR constraints bind only at either or both of the extremes. Consider first this latter case in which the IR constraints bind at both extremes of the interval (one of which may be $\underline{\theta}$ or $\overline{\theta}$ or they both may be $\underline{\theta}$ and $\overline{\theta}$). Denote such an interval by $[\theta', \theta'']$. By construction, the IR constraints only bind at θ' and θ'', so the cumulative multiplier $\gamma_\tau(\theta) = \gamma_\tau \in (0, 1)$ is constant on (θ', θ''). Note that

$$\int_{\theta'}^{\theta''} \overline{u}'(x, \tau)dx = \overline{u}(\theta'', \tau) - \overline{u}(\theta', \tau) = u_\tau(\theta'') - u_\tau(\theta') = \int_{\theta'}^{\theta''} u'_\tau(x)dx = \int_{\theta'}^{\theta''} \nu(q_\tau(x))dx$$
$$\geq \int_{\theta'}^{\theta''} \overline{u}'(x)dx = \overline{u}(\theta'') - \overline{u}(\theta') = u(\theta'') - u(\theta') = \int_{\theta'}^{\theta''} u'(x)dx = \int_{\theta'}^{\theta''} \nu(q(x))dx. \tag{68}$$

Consider the first line in (68). The second equality holds since the IR constraints of types θ' and θ'' bind before and after the transfer by the assumption of the case, and the fourth equality follows from local incentive compatibility. Consider now the second line in (68). The inequality holds because $\overline{u}'(\theta, \tau) \geq \overline{u}'(\theta)$ for each type, as established, and the equalities hold by the same argument that proves that the corresponding equalities in the first line hold. Equation (68) then implies that

$$\int_{\theta'}^{\theta''} \nu(q_\tau(x))dx \geq \int_{\theta'}^{\theta''} \nu(q(x))dx. \tag{69}$$

Since $\nu(\cdot)$ is positive and increasing, (69) implies that $l(\gamma_\tau, \theta) = q_\tau(\theta) \geq q(\theta) = l(\gamma, \theta)$ for a set of types in $[\theta', \theta'']$ with positive measure—recall the definition of $l(\cdot, \cdot)$ from assumption (PS). Since, as argued, $l(\cdot, \theta)$ decreases with γ, the fact that $q_\tau(\theta) \geq q(\theta)$ yields that $\gamma_\tau \leq \gamma$ on this subset of $[\theta', \theta'']$. But given that the multiplier is constant in the interior of $[\theta', \theta'']$ by construction, it must be that $\gamma_\tau \leq \gamma$ on the entire interval (θ', θ''). Therefore, $q_\tau(\theta) \geq q(\theta)$ for all types in (θ', θ''). By local incentive compatibility, then, $T'_\tau(q_\tau(\theta)) = \theta\nu'(q_\tau(\theta)) \leq \theta\nu'(q(\theta)) = T'(q(\theta))$ for all types in (θ', θ'') since $\nu'(\cdot)$ is decreasing. Finally, $q_\tau(\theta) \geq q(\theta)$ and $T'_\tau(q_\tau(\theta)) \leq T'(q(\theta))$ for θ' and θ'' follow immediately from (65), given that the IR constraints bind for θ' and θ'' by the assumption of the case, and so $q_\tau(\theta) = \overline{q}_\tau(\theta)$ and $q(\theta) = \overline{q}(\theta)$.

Consider next the case in which the IR constraints bind only at one of the extremes of the interval. As argued, these last intervals to consider are of the form $[\underline{\theta}, \theta''']$, where the IR constraints bind only at θ''', and $[\theta'''', \overline{\theta}]$, where the IR constraints bind only at θ''''. But then $\gamma_\tau = 0$ for all types in $[\underline{\theta}, \theta''')$ and $\gamma_\tau = 1$ for all types in $[\theta'''', \overline{\theta}]$. Since the binding pattern of the IR constraints is the same before and after the transfer, it must be that $\gamma = \gamma_\tau$, and so $q_\tau(\theta) = q(\theta)$ and $T'_\tau(q_\tau(\theta)) = T'(q(\theta))$ for all types in $[\underline{\theta}, \theta''')$ and $[\theta'''', \overline{\theta}]$. (A similar argument holds when $\gamma_\tau = 0$ or $\gamma_\tau = 1$ for all types.) For θ''', the same argument as the one leading to (65) applies, and so $q_\tau(\theta) \geq q(\theta)$ and $T'_\tau(q_\tau(\theta)) \leq T'(q(\theta))$.

Under both Case 1 and Case 2, given that $F(\theta) = G(q) = G_\tau(q_\tau)$ and $q_\tau(\theta) \geq q(\theta)$ for each type, it follows that the distribution of quantities after the transfer is introduced first-order stochastically dominates the one before the transfer is introduced. Since, as argued, the transfer amounts to a reduction in consumers' reservation utilities in that $\overline{u}(\theta, \tau) \leq \overline{u}(\theta)$, the menu $\{t(\theta), q(\theta)\}$ is still implementable, so profits must weakly increase. Moreover, the offered quantity (weakly) increases for each type, and so the cost of producing each type's quantity is higher after the transfer. Hence, $T(q(\theta))$ must (weakly) increase for each type.

Part II: Consider now the more general case in which the binding pattern of the IR constraints is not the same before and after the transfer. Pick an interval of consumer types whose IR constraints bind before and after the transfer. Then, the same argument as that leading to (65) proves the desired result. Consider now an interval of consumer types with binding IR constraints only at the extremes of such an interval before and after the transfer. Then, the same argument as that leading to (68) and (69) and the discussion following them establish the desired claim. □

Proof of Corollary 2: Recall that $T'(q)=\theta(q)\nu'(q)$ and $T'_\tau(q_\tau)=\theta_\tau(q_\tau)\nu'(q_\tau)$ by local incentive compatibility before and after the transfer, and that $q=q(\theta)$ and $q_\tau=\tilde{q}_\tau(q)=q_\tau(\theta)$. Differentiating these local incentive compatibility conditions gives

$$T''_\tau(q_\tau(\theta))=\frac{1}{q'_\tau(\theta)}\nu'(q_\tau(\theta))+\theta\nu''(q_\tau(\theta)) \text{ and } T''(q(\theta))=\frac{1}{q'(\theta)}\nu'(q(\theta))+\theta\nu''(q(\theta)). \tag{70}$$

Now, $F(\theta)=G_\tau(q_\tau(\theta))$ and $F(\theta)=G(q(\theta))$, respectively, imply that $1/q'_\tau(\theta)=g_\tau(q_\tau)/f(\theta)$ and $1/q'(\theta)=g(q)/f(\theta)$. Then, $q'(\theta)\leq q'_\tau(\theta)$ if, and only if, $g_\tau(q_\tau)\leq g(q)$, since

$$\frac{1}{q'_\tau(\theta)}=\frac{g_\tau(q_\tau)}{f(\theta)}\leq\frac{1}{q'(\theta)}=\frac{g(q)}{f(\theta)}\Leftrightarrow g_\tau(q_\tau)\leq g(q).$$

Suppose that $q_\tau(\theta)\geq q(\theta)$ for each consumer type. Thus, $G_\tau(q_\tau)\leq G(q)$. Since $G_\tau(\cdot)$ first-order stochastically dominates $G(\cdot)$, it follows that $g_\tau(q_\tau)\leq g(q)$ up to a certain type $\theta_{\max}\leq\overline{\theta}$. Hence, $q'(\theta)\leq q'_\tau(\theta)$ for any $\theta\leq\theta_{\max}$. Also, $q_\tau(\theta)\geq q(\theta)$, by assumption, for all types implies that $\nu'(q_\tau(\theta))\leq\nu'(q(\theta))$ if $\nu''(\cdot)\leq 0$, which is a maintained assumption, and $\nu''(q_\tau(\theta))\leq\nu''(q(\theta))$ if $\nu'''(\cdot)\leq 0$, which holds by assumption. Thus, by inspecting (70) it is immediate that $T''_\tau(\tilde{q}_\tau(q))\leq T''(q)$ for all $\theta\leq\theta_{\max}$ or, equivalently, for all percentiles of quantities up to the percentile $p_{\max}$ corresponding to the quantity $q(\theta_{\max})$ before the transfer and to the quantity $q_\tau(\theta_{\max})$ after the transfer, as desired. □

No Resale: We interpret the fact that it may be difficult for consumers to engage in the type of contracts that would sustain resale as capturing imperfections in contracting among consumers analogous to those we maintain among sellers. The argument is simple. When consumers' characteristics are observable but not contractible to consumers too, the problem that a coalition of consumers would face at the resale stage is a constrained version of the one that the seller faces. In particular, the coalition's problem of maximizing consumers' utility in excess of the utility each type achieves by purchasing from the seller, that is, $s(\theta,q(\theta))-u(\theta)$, would be identical to the seller's problem given that $s(\theta,q(\theta))-u(\theta)=t(\theta)-c(q(\theta))$, except for the additional constraint of linear pricing, if only linear prices were enforceable by the coalition. If enforcement or transaction costs, due, for instance, to commuting across villages, were of the order of $s(\theta,q(\theta))-u(\theta)$ or higher for each type θ, then the coalition could not achieve higher utility for any member than the utility each member obtains by trading with the seller. See Ligon et al. (2002) for evidence on contracting imperfections in developing countries. □

Example 1 (Nonlinear vs. Linear Pricing): Suppose the base utility function, $\nu(q)$, is a three-parameter HARA function with $\nu(q)=(1-d)[aq/(1-d)+b]^d/d$, $a>0$, $aq/(1-d)+b>0$, and $0<d<1$. Denote by $u_s(\theta)$ the utility of a type θ consumer under the standard model. With a uniform type distribution on $[\underline{\theta},\overline{\theta}]$, $\overline{u}(\underline{\theta})=0$, $a=c=1$, $b=0$, and $d=1/2$, it follows $u_s(\theta)\geq u_m(\theta)$ if, and only if, $(2\theta-\overline{\theta})^2-(2\underline{\theta}-\overline{\theta})^2\geq\theta^2$. When $\underline{\theta}=1$ and $\overline{\theta}=2$, this expression reduces to $3\theta^2-8\theta+4\geq 0$, a polynomial with roots $\theta=2/3$ and $\theta=2$. Thus, $u_m(\theta)\geq u_s(\theta)$ for *all* consumer types. Consider now the highly convex case of the augmented model with $\gamma=1/2$. In this case, $u(\theta)\geq u_m(\theta)$ if, and only if, $3\theta^2-6\theta+2\geq 0$. So, $u(\theta)\geq u_m(\theta)$ for $\theta\geq 1.58$. Also, all such types demand quantities above first best. Thus, not only $u(\theta)\geq u_s(\theta)$, as implied by Proposition 4, but also more than half of the consumers prefer nonlinear to linear pricing under the augmented model, whereas all consumers prefer linear to nonlinear pricing under the standard model. □

Proof of Proposition 8: Rewrite the seller's first-order condition as

$$\frac{1}{T'(q)}=\frac{1}{c'(Q)}+\frac{[F(\theta)-\gamma(\theta)]}{c'(Q)\theta f(\theta)}=\frac{1}{c'(Q)}+\frac{[G(q)-\gamma(\theta(q))]\theta'(q)}{c'(Q)g(q)\theta(q)}=\frac{1}{c'(Q)}+\frac{[G(q)-\gamma(\theta(q))]\varphi(q)}{c'(Q)g(q)} \tag{71}$$

with $\varphi(q)\equiv\partial\log(\theta(q))/\partial q=\theta'(q)/\theta(q)$ or, equivalently,

$$\frac{g(q)}{\varphi(q)}\left[\frac{c'(Q)}{T'(q)}-1\right]=G(q)-\psi(q),$$

where $\psi(q) \equiv \gamma(\theta(q))$. By taking derivatives of each side of this expression, we obtain

$$\frac{\partial\{g(q)c'(Q)/[\varphi(q)T'(q)]\}}{\partial q} - \frac{\partial[g(q)/\varphi(q)]}{\partial q} = g(q) - \psi'(q).$$

Integrating these expressions from $\underline{q}$ to $\overline{q}$ gives

$$\int_{\underline{q}}^{\overline{q}} \frac{\partial\{g(x)c'(Q)/[\varphi(x)T'(x)]\}}{\partial x} dx - \int_{\underline{q}}^{\overline{q}} \frac{\partial[g(x)/\varphi(x)]}{\partial x} dx = \int_{\underline{q}}^{\overline{q}} g(x)dx - \int_{\underline{q}}^{\overline{q}} \psi'(x)dx = 0,$$

where the last equality follows from the fact that $\int_{\underline{q}}^{\overline{q}} g(x)dx = \int_{\underline{q}}^{\overline{q}} \psi'(x)dx = 1$, so that

$$\frac{g(\overline{q})c'(Q)}{\varphi(\overline{q})T'(\overline{q})} - \frac{g(\underline{q})c'(Q)}{\varphi(\underline{q})T'(\underline{q})} - \frac{g(\overline{q})}{\varphi(\overline{q})} + \frac{g(\underline{q})}{\varphi(\underline{q})} = 0,$$

which implies

$$c'(Q) = \left[g(\overline{q}) - g(\underline{q})\frac{\varphi(\overline{q})}{\varphi(\underline{q})}\right] / \left[\frac{g(\overline{q})}{T'(\overline{q})} - \frac{g(\underline{q})}{T'(\underline{q})}\frac{\varphi(\overline{q})}{\varphi(\underline{q})}\right].$$

Since $g(q)$ and $T'(q)$ are identified, it follows that $c'(Q)$ is identified up to $\varphi(\overline{q})/\varphi(\underline{q})$. The rest of the proposition is proved in the main text. □

Derivation of Reduced Form in (18): The seller's first-order condition can be rewritten as

$$\frac{T'(q) - c'(Q)}{T'(q)} = \frac{\gamma(\theta) - F(\theta)}{\theta f(\theta)} = \frac{\theta'(q)}{\theta(q)}\left[\frac{\gamma(\theta(q)) - G(q)}{g(q)}\right] \Leftrightarrow \frac{c'(Q)}{T'(q)} = 1 + \frac{\theta'(q)}{\theta(q)}\left[\frac{G(q)}{g(q)} - \frac{\gamma(\theta(q))}{g(q)}\right],$$

by using $F(\theta) = G(q)$, $f(\theta) = g(q)q'(\theta)$, and $\theta(q)q'(\theta) = \theta(q)/\theta'(q)$. Further manipulation yields

$$\log\left(\frac{c'(Q)}{T'(q)}\right) \approx \frac{\theta'(q)}{\theta(q)}\left[\frac{G(q)}{g(q)} - \frac{\gamma(\theta(q))}{g(q)}\right] \Leftrightarrow \log\left(\frac{p(q)}{\frac{c'(Q)}{t_1}}\right) \approx \frac{\theta'(q)}{\theta(q)}\left[\frac{\gamma(\theta(q)) - 1}{g(q)} + \frac{1 - G(q)}{g(q)}\right]$$
$$\Leftrightarrow \log[p(q)] \approx \log\left[\frac{c'(Q)}{t_1}\right] - \frac{\theta'(q)}{\theta(q)}\left[\frac{1 - \gamma(\theta(q))}{g(q)}\right] + \frac{\theta'(q)}{\theta(q)}\left[\frac{1 - G(q)}{g(q)}\right], \tag{72}$$

using the fact that $T'(q) = t_1 p(q)$ given our log-linear specification for $T(q)$. Note that $[1 - G(q)]/g(q)$ is the inverse of the hazard rate of the distribution quantities. Letting $\gamma(\theta(q)) = \psi(q)$, (72) becomes

$$\log[p(q)] \approx \log\left[\frac{c'(Q)}{t_1}\right] - \frac{\theta'(e^{\log(q)})}{\theta(e^{\log(q)})}\left[\frac{1 - \psi(e^{\log(q)})}{g(e^{\log(q)})}\right] + \frac{\theta'(e^{\log(q)})}{\theta(e^{\log(q)})}\left[\frac{1 - G(q)}{g(q)}\right].$$

We can interpret the right side of this expression as a function of $\log(q)$ and $[1 - G(q)/g(q)]$, which we term $f(\log(q), [1 - G(q)]/g(q))$. A second-order Taylor expansion of $f(\cdot)$ in a neighborhood of $(\log(q), [1 - G(q)/g(q)]) = (a, b)$ gives

$$\log[p(q)] \approx f(a,b) + f_1(a,b)(\log(q) - a) + f_2(a,b)\left[\frac{1 - G(q)}{g(q)} - b\right]$$
$$+ \frac{1}{2}\left\{f_{11}(a,b)[\log(q) - a]^2 + 2f_{12}(a,b)[\log(q) - a]\left[\frac{1 - G(q)}{g(q)} - b\right] + f_{22}(a,b)\left[\frac{1 - G(q)}{g(q)} - b\right]^2\right\},$$

and so

$$\log[p(q)] \approx \underbrace{f(a,b) - af_1(a,b) - bf_2(a,b) + \frac{a^2}{2}f_{11}(a,b) + \frac{b^2}{2}f_{22}(a,b) + abf_{12}(a,b)}_{\beta_0}$$

$$+ \underbrace{[f_1(a,b) - af_{11}(a,b) - bf_{12}(a,b)]}_{\beta_1}\log(q) + \underbrace{[f_2(a,b) - af_{12}(a,b) - bf_{22}(a,b)]}_{\beta_2}\left[\frac{1-G(q)}{g(q)}\right]$$

$$+ \underbrace{f_{12}(a,b)}_{\beta_3}\log(q)\left[\frac{1-G(q)}{g(q)}\right] + \underbrace{\frac{1}{2}f_{11}(a,b)}_{\beta_4}\log(q)^2 + \underbrace{\frac{1}{2}f_{22}(a,b)}_{\beta_5}\left[\frac{1-G(q)}{g(q)}\right]^2,$$

which can be equivalently expressed as

$$\log[p(q)] \approx \beta_0 + \beta_1\log(q) + \beta_2\left[\frac{1-G(q)}{g(q)}\right] + \beta_3\log(q)\left[\frac{1-G(q)}{g(q)}\right] + \beta_4\log(q)^2 + \beta_5\left[\frac{1-G(q)}{g(q)}\right]^2. \quad \square$$

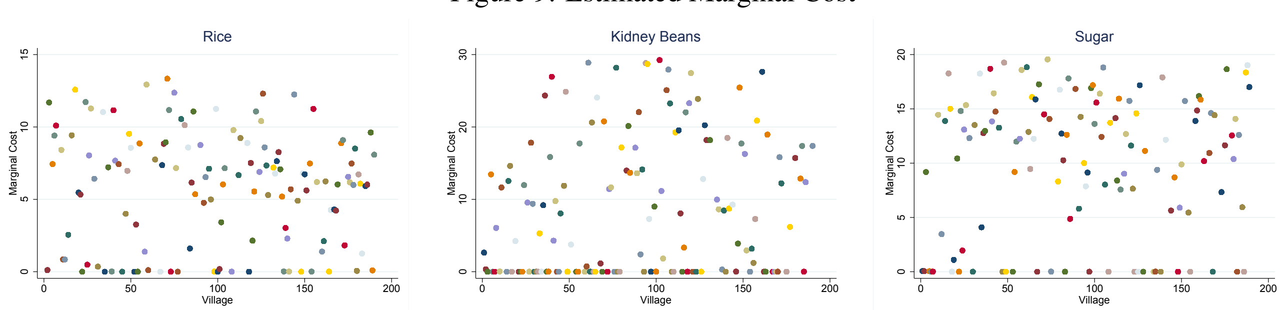
B Appendix

B.1 Omitted Estimation Results

We present here estimation results omitted from the main text for brevity.

Marginal Cost Estimates. Figure 9 reports the estimated marginal cost at the total quantity provided of each good in each estimated village according to our quadratic specification of the multiplier function; we truncated the top 5% of the estimates for readability. The estimates for the linear specification are very similar; see Section B.3. The mean of the estimated marginal cost across villages defined as municipalities is 6.342 pesos for rice with a standard deviation of 4.120, 10.839 pesos for kidney beans with a standard deviation of 20.647, and 9.834 pesos for sugar with a standard deviation of 7.837. Although estimated marginal cost varies noticeably across villages for each good, its range is similar for rice and sugar, from close to 0 pesos to about 15 pesos, and its mean comparable across kidney beans and sugar. Villages are very dispersed and isolated, so some variability in estimated marginal cost across villages is to be expected.

Figure 9: Estimated Marginal Cost



Predicted Multiplier Schedule. We perform significance tests on the parameters of the multiplier function $\gamma(\theta(q))$ in each village and for each good so as to determine whether the multiplier is ever constant and construct the predicted multiplier for each quantity accordingly. We illustrate our procedure only in the quadratic case, since the linear case is a special case of it. Consider village v and good j. First, we test whether the three parameters, γ_{vj0}, γ_{vj1}, and γ_{vj2}, are individually significant. If they are, then we use all of them to construct the predicted $\gamma(\theta(q))$. If only γ_{vj0} is significant, then we set all other parameters to zero and next test whether the implied multiplier is equal to 0 or 1. If we reject these hypotheses, then we keep the original value of γ_{vj0}; otherwise, we set the multiplier to 0 or 1 depending on the outcome of the test. If only γ_{vj1} is significant, then we set all other parameters to zero. If only γ_{vj2} is significant, then we set all other parameters to zero. If only γ_{vj0} and γ_{vj1} are significant, then we set γ_{vj2} to zero. If only γ_{vj0}

and γ_{vj2} are significant, then we set γ_{vj1} to zero. If only γ_{vj1} and γ_{vj2} are significant, then we set γ_{vj0} to zero. If no parameter is significant, then we set all of them to zero so that $\gamma(\theta(q))$ equals 0.5 at all quantities.

Statistics on Estimates. Tables 3 to 5 report selected percentiles of the distribution of the t-statistics of the estimates of $c'(Q)$, $\gamma(\theta(q))$, $\theta(q)$, $\nu'(q)$, and $f(\theta)$ across villages. These statistics are meant to illustrate the overall precision of our estimates. We then calculate selected percentiles of the distribution across village-level quantities of the t-statistics of the estimates of $\gamma(\theta(q))$, $\theta(q)$, $\nu'(q)$, and $f(\theta)$ for each village. Tables 6 to 8 report the quartiles of the distribution of these selected percentiles across villages. These statistics are meant to show the variability across villages of the precision of the estimates of $\gamma(\theta(q))$, $\theta(q)$, $\nu'(q)$, and $f(\theta)$ at the different quantities consumed of a good in a village. As apparent from these tables, the model's primitives as well as the multiplier $\gamma(\theta(q))$ are fairly precisely estimated in most villages and for most quantities.

Table 3: Percentiles of t-Statistics across Quantities and Villages for Rice

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.029	0.465	1.569	5.737	13.606	24.958	46.042	57.108	123.646
$\gamma(\theta(q))$	1.177	3.891	6.062	22.413	112.474	1.653×10^3	3.781×10^4	3.630×10^5	7.131×10^7
$\theta(q)$	0.045	0.113	0.363	0.967	2.735	7.174	22.840	46.671	138.463
$\nu'(q)$	-109.942	-26.071	-13.560	-3.979	-0.980	-0.008	10.468	32.654	189.972
$f(\theta)$	1.118	1.118	1.118	3.665	9.671	15.163	24.950	29.262	50.702

Table 4: Percentiles of t-Statistics across Quantities and Villages for Kidney Beans

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.015	0.047	0.165	2.363	7.364	15.099	26.034	35.793	60.875
$\gamma(\theta(q))$	2.585	5.642	10.147	31.358	156.134	492.064	3.603×10^3	2.231×10^4	1.627×10^5
$\theta(q)$	9.029×10^{-18}	0.135	0.519	1.458	3.513	7.078	11.241	15.223	23.309
$\nu'(q)$	-15.049	-9.306	-7.152	-3.434	-1.202	0.615	13.109	30.124	80.267
$f(\theta)$	1.118	1.118	1.581	5.244	11.565	18.473	24.239	27.613	36.652

Table 5: Percentiles of t-Statistics across Quantities and Villages for Sugar

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.004	0.012	0.081	3.668	11.742	30.069	72.451	129.396	274.517
$\gamma(\theta(q))$	1.166	4.322	8.200	32.756	170.466	575.413	3.650×10^3	1.795×10^4	1.749×10^6
$\theta(q)$	0.002	0.225	0.414	1.127	2.683	5.862	12.785	18.854	49.134
$\nu'(q)$	-28.371	-12.893	-8.537	-4.123	-1.345	-0.326	2.810	9.445	51.105
$f(\theta)$	1.118	1.118	1.118	6.800	12.942	18.874	24.520	27.407	35.777

Table 6: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Rice

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	0.940	2.901	4.969	8.183	23.735	114.980	3.114×10^3	2.894×10^4	6.335×10^5
	p_{50}	1.502	10.257	15.798	32.789	84.607	635.232	9.581×10^3	1.922×10^5	1.776×10^6
	p_{75}	8.125	24.850	44.102	125.823	397.046	3.300×10^3	7.418×10^4	2.370×10^6	2.194×10^8
$\theta(q)$	p_{25}	0.024	0.087	0.281	0.704	1.725	4.215	14.759	38.274	66.776
	p_{50}	0.082	0.155	0.495	1.171	2.541	6.509	18.461	58.424	99.308
	p_{75}	0.148	0.543	0.833	1.805	4.076	9.827	28.787	77.431	119.838
$\nu'(q)$	p_{25}	−109.942	−49.192	−22.981	−6.238	−2.203	−0.848	−0.243	−0.125	1.199
	p_{50}	−83.125	−33.848	−13.560	−3.826	−0.928	−0.101	1.537	4.799	63.914
	p_{75}	−28.180	−16.361	−5.874	−1.777	−0.092	3.762	24.882	65.786	188.850
$f(\theta)$	p_{25}	1.118	1.118	1.118	2.168	4.330	9.093	15.330	28.657	50.279
	p_{50}	2.739	5.000	6.120	8.281	11.307	15.163	27.161	31.031	50.727
	p_{75}	5.123	6.703	7.583	9.269	12.500	17.746	27.166	31.055	50.727

Table 7: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Kidney Beans

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	2.222	4.049	8.732	12.704	34.369	112.622	231.876	295.473	908.156
	p_{50}	4.337	12.143	21.240	41.340	107.275	231.221	436.778	626.987	3.432×10^3
	p_{75}	20.332	65.232	115.150	172.986	411.486	1.300×10^3	2.838×10^3	6.372×10^3	4.907×10^4
$\theta(q)$	p_{25}	0.001	0.120	0.384	0.927	2.628	5.160	8.517	10.301	15.412
	p_{50}	0.024	0.271	0.576	1.458	3.547	6.453	9.597	13.718	21.626
	p_{75}	0.024	0.438	0.907	2.165	4.421	8.372	13.787	17.482	24.552
$\nu'(q)$	p_{25}	−15.421	−12.178	−9.057	−5.783	−2.612	−1.101	−0.306	−0.047	2.284
	p_{50}	−10.922	−7.841	−5.372	−2.967	−1.324	−0.185	1.481	2.389	35.157
	p_{75}	−7.341	−4.222	−3.387	−1.170	0.429	5.723	11.765	21.521	66.214
$f(\theta)$	p_{25}	1.118	1.581	1.809	3.162	5.477	9.417	13.509	16.508	31.605
	p_{50}	4.183	6.349	7.804	10.094	15.026	20.520	27.060	32.365	36.646
	p_{75}	7.045	8.657	9.421	12.298	16.771	21.700	27.458	32.365	36.663

Table 8: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Sugar

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	1.248	2.004	5.879	12.244	44.027	191.167	424.775	2.054×10^3	6.874×10^5
	p_{50}	3.344	12.293	20.236	38.504	119.964	289.418	888.764	6.885×10^3	7.345×10^5
	p_{75}	14.400	31.077	39.699	124.092	345.134	886.617	3.359×10^3	6.000×10^4	2.943×10^6
$\theta(q)$	p_{25}	7.700×10^{-9}	0.144	0.281	0.906	2.001	3.989	8.619	10.142	22.735
	p_{50}	0.082	0.303	0.797	1.109	2.721	6.396	11.350	16.743	31.355
	p_{75}	0.111	0.402	0.978	1.869	3.790	8.155	17.517	25.216	113.818
$\nu'(q)$	p_{25}	−95.314	−16.733	−12.057	−6.204	−2.784	−1.144	−0.564	−0.326	−0.094
	p_{50}	−21.357	−9.803	−7.400	−3.830	−1.654	−0.730	−0.208	0.014	0.810
	p_{75}	−11.786	−5.866	−4.410	−1.864	−0.635	0.474	3.811	8.184	70.169
$f(\theta)$	p_{25}	1.118	1.118	1.350	3.873	6.982	10.062	13.229	19.170	23.372
	p_{50}	3.741	7.746	9.152	12.349	15.652	20.767	25.125	26.972	34.424
	p_{75}	6.801	8.056	9.485	13.038	17.045	22.220	27.407	32.215	36.552

B.2 Details on Counterfactuals

In computing counterfactuals for each good and village, we exclude observations for which any of the objects of interest in a village, that is, either the model's primitives, $c'(Q)$, $\nu'(q)$, and $\theta(q)$, or objects used in their estimation, $G(q)$, $g(q)$, $T(q)$, $T'(q)$, and $\gamma(\theta(q))$, were not successfully estimated. We also drop observations for which we estimate an implausibly large increase in $\theta(q)$, that is, whenever $\theta(q_i)$ for quantity q_i is more than $1,000$ times larger than its values for the previous quantity, $\theta(q_{i-1})$. If this is the case, then we drop such a quantity observation and all subsequent observations in the village considered. In terms of notation, we denote by $\widehat{\theta}_i = \widehat{\theta}(q_i)$ the estimated type consuming quantity q_i and by $\widehat{f}$ the estimate of the object f. For simplicity, we focus here on a particular good and village so that the associated subscripts are omitted.

Base Utility Function. For the first quantity in each village, we set

$$\widehat{\nu}(q_1) = \begin{cases} q_i\widehat{\nu}'(q_1), & \text{if } q_1\widehat{\nu}'(q_1) - T(q_1)/\widehat{\theta}(q_1) \geq 0 \\ T(q_1)/\widehat{\theta}(q_1), & \text{otherwise} \end{cases}.$$

This approach guarantees that base utility at the lowest quantity is nonnegative. For the remaining quantities, we compute base utility using a standard Riemann sum approximation as

$$\widehat{\nu}(q_i) = q_1\widehat{\nu'}(q_1) + \sum_{j=1}^{i-1}(q_{j+1} - q_j)\widehat{\nu'}(q_{j+1}).$$

Given $\widehat{\nu}(q_i)$ evaluated at the observed quantities $\{q_i\}$, we compute $\widehat{\nu}(q)$ at any other quantity q as the predicted value from a regression of $\widehat{\nu}(q_i)$ on $\log(q_i)$, that is, $\widehat{\nu}(q) = \widehat{\beta}_0 + \widehat{\beta}_1 \log(q)$.

Reservation Utility. Recall that $\widehat{\gamma}(\widehat{\theta}_i) = \widehat{\gamma}(\widehat{\theta}_i(q_i))$ is estimated as a function of quantity and weakly increases with $\widehat{\theta}_i$. We set $\overline{u}(\widehat{\theta}_1) = \widehat{CS}_{np}(\widehat{\theta}_1)$, where $\widehat{CS}_{np}(\widehat{\theta}_1)$ is the estimated consumer surplus of type $\widehat{\theta}_1$ under nonlinear pricing. For the remaining consumer types, we consider two cases:

Case 1: When $\widehat{\gamma}(\widehat{\theta}_i) \geq \widehat{\gamma}(\widehat{\theta}_{i-1})$, we test whether $\widehat{\gamma}(\widehat{\theta}_i)$ is statistically different from $\widehat{\gamma}(\widehat{\theta}_{i-1})$. To do so, we use a critical value of the Student's t-distribution for a 1% confidence level. If $\widehat{\gamma}(\widehat{\theta}_i)$ is statistically different from $\widehat{\gamma}(\widehat{\theta}_{i-1})$, then we set $\overline{u}(\widehat{\theta}_i) = \widehat{CS}_{np}(\widehat{\theta}_i)$. If, instead, $\widehat{\gamma}(\widehat{\theta}_i)$ is not statistically different from $\widehat{\gamma}(\widehat{\theta}_{i-1})$, then we set $\overline{u}(\widehat{\theta}_i) = \overline{u}(\widehat{\theta}_{i-1})$.

Case 2: When $\widehat{\gamma}(\widehat{\theta}_i)$ is constant for each type, that is, $\widehat{\gamma}(\widehat{\theta}_i) = \widehat{\gamma}(\widehat{\theta}_{i-1})$ for each $\widehat{\theta}_{i-1}$ and $\widehat{\theta}_i$, we consider two possibilities, referred to in the main text as the low reservation utility case and the high reservation utility case. In the low reservation utility case, we set $\overline{u}(\widehat{\theta}_i) = \widehat{CS}_{np}(\widehat{\theta}_1)$ except for the highest type, $\overline{u}(\widehat{\theta}_N)$, for which we set $\overline{u}(\widehat{\theta}_N) = \widehat{CS}_{np}(\widehat{\theta}_N)$, where q_N is the largest quantity purchased in a village. In the high reservation utility case, we set $\overline{u}(\widehat{\theta}_i) = \widehat{CS}_{np}(\widehat{\theta}_i)$ for each $\widehat{\theta}_i$.

First-Best Counterfactual. When comparing consumer, producer, and social surplus under nonlinear and first-best pricing, we start by computing the quantities demanded under first best. To this purpose, we define a grid $\mathbf{q_{pc}} = (q_1, \ldots, q_{\max})$ of 10,000 equidistant points for candidate quantities. Given that the first-best price p_{pc} equals the marginal cost, each type $\widehat{\theta}_i$ solves the problem $\max_{q \in \mathbf{q_{pc}}}[\widehat{\theta}_i\widehat{\nu}(q) - p_{pc}q]$. Then, the quantity demanded by type $\widehat{\theta}_i$ at price p_{pc} is

$$q_{pc}(\widehat{\theta}_i) = \begin{cases} \widetilde{q}(\widehat{\theta}_i), & \text{if } \widehat{\theta}_i\widehat{\nu}(\widetilde{q}(\widehat{\theta}_i)) - p_{pc}\widetilde{q}(\widehat{\theta}_i) \geq \overline{u}(\widehat{\theta}_i) \\ 0, & \text{otherwise} \end{cases},$$

with $\overline{u}(\widehat{\theta}_i)$ is determined as just explained. Then, consumer surplus is

$$\widehat{CS}_{pc} = \sum_{i=1}^{N-1} r(\widehat{\theta}_i) \max\{\widehat{\theta}_i\widehat{\nu}(q_{pc}(\widehat{\theta}_i)) - p_{pc}q_{pc}(\widehat{\theta}_i), \overline{u}(\widehat{\theta}_i)\},$$

where $r(\widehat{\theta}_i)$ is the proportion of type $\widehat{\theta}_i$ in the village, computed as $r(\widehat{\theta}_1) = G(q_1)$ and $r(\widehat{\theta}_{i+1}) = G(q_{i+1}) - G(q_i)$ for $i = 1, \ldots, N-1$, and N is the number of consumers in the village. In plotting results, we only consider types with $\widehat{CS}_{np}(\widehat{\theta}_i) \geq 0$ and $\widehat{SS}_{pc}(\widehat{\theta}_i) > \widehat{CS}_{np}(\widehat{\theta}_i) + \widehat{PS}_{np}(\widehat{\theta}_i)$, that is, with nonnegative consumer surplus under nonlinear pricing and positive social surplus under first best; we consider this latter condition as a regularity condition. We trim observations of quantities corresponding to types in the top 2% of the type distribution in a village for the readability of the graphs in the main text.

Linear Pricing Counterfactual. When comparing consumer, producer, and social surplus under nonlinear and linear pricing, we compute a seller's linear price, individual demand, and aggregate demand as follows. Let $\widetilde{\Theta}$ be the set of estimated types $\widehat{\theta}_i = \widehat{\theta}(q_i)$ with $\widehat{CS}_{np}(\widehat{\theta}_i) \geq 0$ and $\widehat{SS}_{pc}(\widehat{\theta}_i) > \widehat{CS}_{np}(\widehat{\theta}_i) + \widehat{PS}_{np}(\widehat{\theta}_i)$, that is, with nonnegative consumer surplus under nonlinear pricing and positive social surplus under first best; as mentioned, we consider this latter condition as a regularity condition. We compute the price p_m^* that maximizes profits under linear pricing as

$$\max_{p_m}[(p_m - c)Q(p_m)], \tag{73}$$

where $Q(p_m)$ is aggregate demand, computed as described next. To solve for p_m, we set a grid of 10,000 equidistant points, $\mathbf{p} = (p_{m1}, \ldots, p_{mP})$, and for each $\widehat{\theta}_i$ we compute the schedule

$$q((p_{m1}, \ldots, p_{mP}), \widehat{\theta}_i) = (q(p_{m1}, \widehat{\theta}_i), \ldots, q(p_{mP}, \widehat{\theta}_i)).$$

To do so, given a grid $\mathbf{q} = (q_1, \ldots, q_{\max})$ of 10,000 equidistant points for candidate quantities demanded, we determine the quantity chosen by type $\widehat{\theta}_i$ for each possible price p_{mp} by solving $\tilde{q}(p_{mp}, \widehat{\theta}_i) = \arg\max_{q \in \mathbf{q}}[\widehat{\theta}_i \widehat{\nu}(q) - p_{mp} q]$. Then, the quantity demanded by type $\widehat{\theta}_i$ at price p_{mp} is

$$q(p_{mp}, \widehat{\theta}_i) = \begin{cases} \widetilde{q}(p_{mp}, \widehat{\theta}_i), \text{ if } \widehat{\theta}_i \widehat{\nu}(\widetilde{q}(p_{mp}, \widehat{\theta}_i)) - p_{mp} \widetilde{q}(p_{mp}, \widehat{\theta}_i) \geq \overline{u}(\widehat{\theta}_i) \\ 0, \text{ otherwise} \end{cases},$$

with $\overline{u}(\widehat{\theta}_i)$ determined as explained above. Aggregate demand for each price p_{mp} is

$$Q(p_{mp}) = \sum_{\widehat{\theta}_i \in \widetilde{\Theta}} r(\widehat{\theta}_i) q(p_{mp}, \widehat{\theta}_i),$$

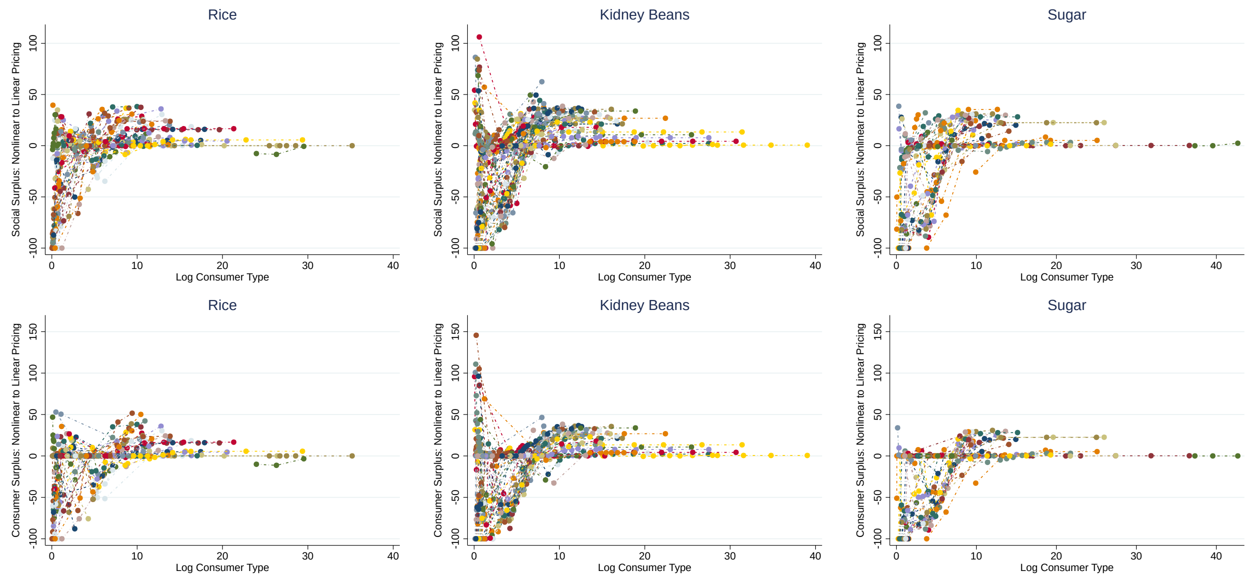
where $r(\widehat{\theta}_1) = G(q_1)$ and $r(\widehat{\theta}_{i+1}) = G(q_{i+1}) - G(q_i)$ for $i = 1, \ldots, N-1$, and N is the number of consumers in the village. We then solve for the price p_m^* such that $(p_m - c)Q(p_m)$ is maximal. We calculate consumer surplus as

$$\widehat{CS}_{lp} = \sum_{\widehat{\theta}_i \in \widetilde{\Theta}} r(\widehat{\theta}_i) \max\{\widehat{\theta}_i \widehat{\nu}(q(p_m^*, \widehat{\theta}_i)) - p_m^* q(p_m^*, \widehat{\theta}_i), \overline{u}(\widehat{\theta}_i)\}.$$

We compute producer surplus as $\widehat{PS}_{lp} = (p_m^* - c)Q(p_m^*)$. In plotting results, we exclude observations in the top 2% of the type distribution in a village for the readability of the graphs in the main text.

Nonlinear vs. Linear Pricing: High Reservation Utility. We present in Figure 10 the results of the linear pricing counterfactual in the high reservation utility case. As in the case of low reservation utility, social surplus at low quantities is higher, and most consumers of the lowest quantities are better off under nonlinear pricing, than under linear pricing. As in the case of low reservation utility, social surplus as well as consumers surplus are lower under nonlinear pricing for consumers of intermediate to large quantities. Intuitively, the key difference between the low and high reservation utility case is the level of the linear price that a seller can charge without inducing any consumer type to drop out of the market. In the low reservation utility case, a seller can charge any given consumer type a much higher price than in the high reservation utility case before that type prefers not to participate. Lower linear prices in the high reservation utility case, compared to the low reservation utility case, benefit consumers with low and high marginal willingness to pay. For some low consumer types, in particular, the combination of high reservation utility and sellers' lower ability to extract surplus under linear pricing implies higher levels of utility under linear pricing than under nonlinear pricing, and so a *reversal* of preferences between nonlinear and linear pricing relative to the low reservation utility case. Nonetheless, consumers who do not participate in the market under linear pricing are still mostly those with low to intermediate types, as in the low reservation utility case. These consumers prefer nonlinear to linear pricing.

Figure 10: Nonlinear vs. Linear Pricing under Augmented Model (High Reservation Utility)



B.3 Robustness Exercises

We assess here the robustness of our empirical findings to specific functional forms and modeling assumptions. We do so by first reestimating the model under a more restrictive specification of the multiplier function, which is assumed now to be a logistic function of quantity with a linear index. We find that the estimates based on this more restricted specification of our model are similar to the ones based on the more general specification discussed in the main text, both qualitatively and quantitatively. We then reestimate the model under the assumption that it reduces to the standard model of Maskin and Riley (1984), that is, by imposing that the multiplier function equals one for each type (quantity) in each village. The estimates implied by the standard model differ from those implied by our model in two main dimensions: the range of the support of consumer types is less wide, and base marginal utility is less convex, according to the standard model than according to our model. As discussed in the main text, though, we reject the standard model in all villages. Finally, we estimate the model's primitives on villages defined at the level of the Mexican locality rather than the Mexican municipality. The estimates we obtain are similar for these two definitions of villages. In the following figures and tables, we refer to the linear (respectively, quadratic) specification of the multiplier function as either the "Linear Specification" (respectively, "Quadratic Specification") or, for brevity, "Linear" (respectively, "Quadratic"). Similar, we refer to kidney beans at times as "Beans" for brevity.

B.3.1 Linear Specification

In Figures 11-14, we plot the estimates of marginal cost, $c'(Q)$, the multiplier on participation (or budget) constraints, $\gamma(\theta(q))$, the distribution of types, as captured by its support, $\theta(q)$, and its reverse hazard rate, $f(\theta)/F(\theta)$, and the base marginal utility function, $\nu'(q)$, for the version of our model in which $\gamma(\theta(q))$ is specified as a logistic function of quantity with a linear index, that is, when we set $\gamma_{vj2} = 0$ for each village v and good j. As is apparent from these figure, all estimates are fairly similar to those reported in the main text, which are obtained under a logistic specification of the multiplier with a quadratic index. As mentioned in the main text, we focus there on the quadratic specification due to the greater number of villages for which we could estimate all primitives, as discussed in Section 5. As before, we drop the top 5% of estimates for readability in the plots of the estimated marginal costs.

Figure 11: Estimated Marginal Cost (Linear Specification)

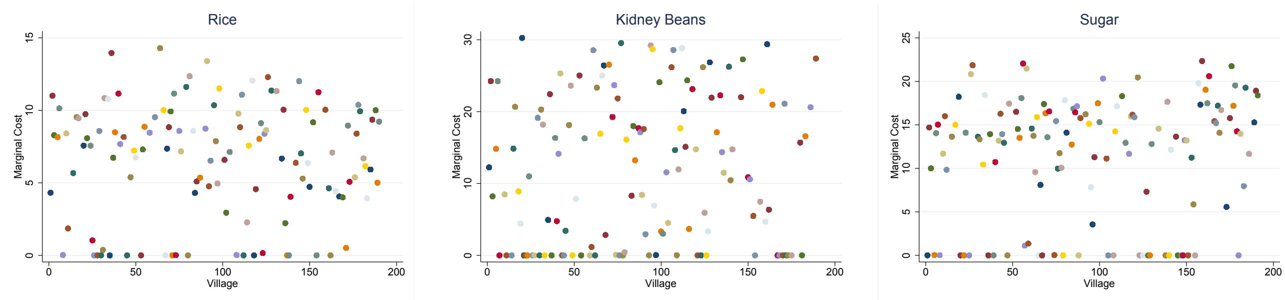


Figure 12: Estimated Multiplier on Participation (or Budget) Constraints (Linear Specification)

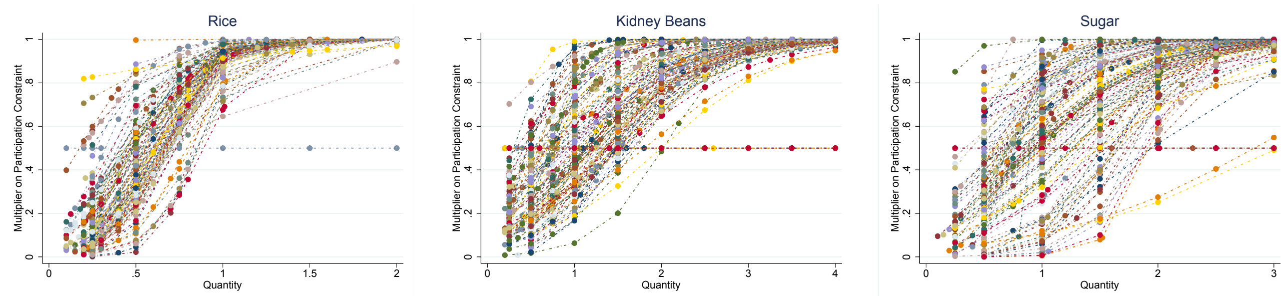


Figure 13: Estimated Distribution of Types (Linear Specification)

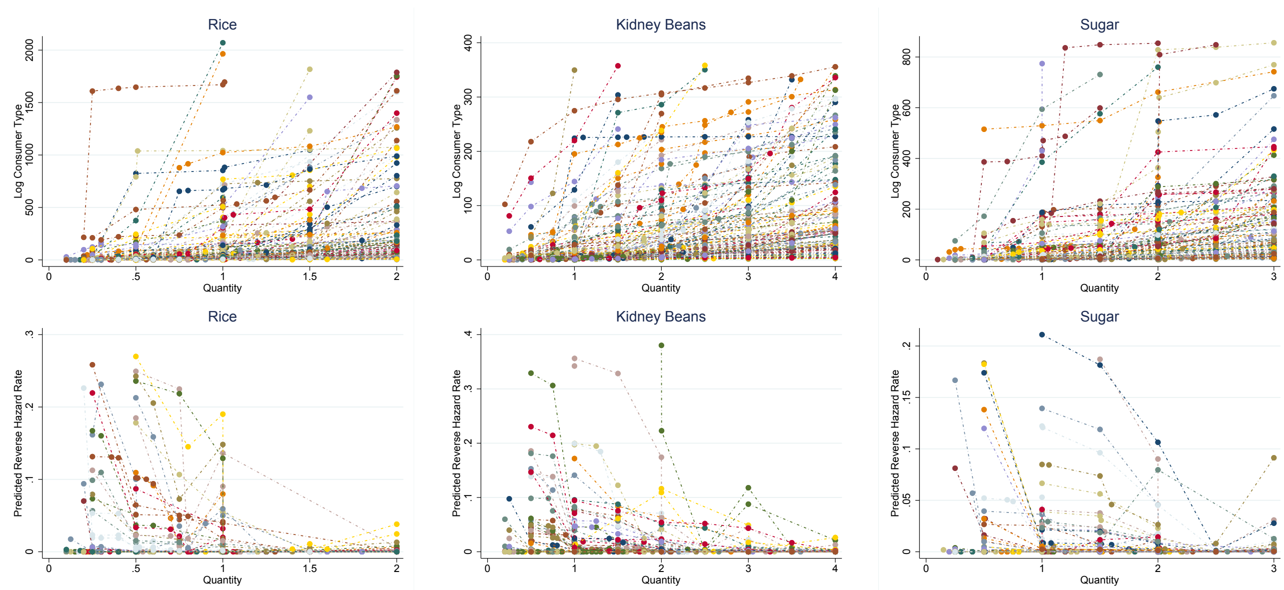


Figure 14: Estimated Base Marginal Utility (Linear Specification)

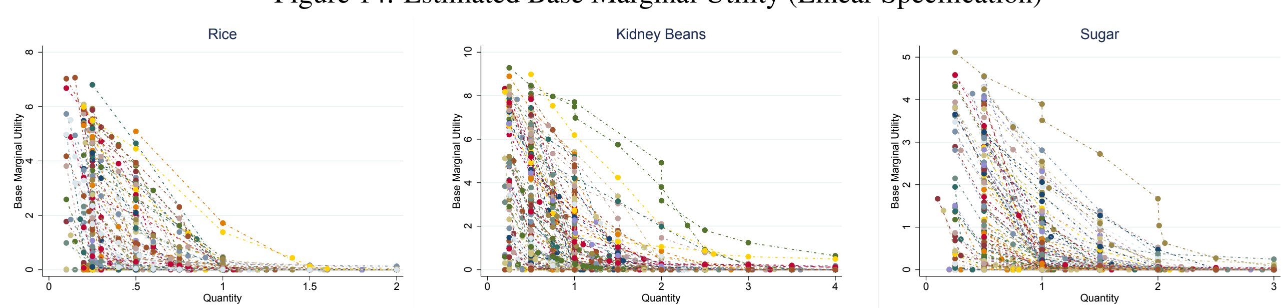


Table 9: Percentiles of t-Statistics across Quantities and Villages for Rice (Linear Specification)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.026	0.179	0.795	6.767	13.553	24.958	42.519	59.439	155.314
$\gamma(\theta(q))$	1.076	6.195	11.401	39.296	244.326	3934.064	8.7×10^4	6.1×10^5	3.0×10^7
$\theta(q)$	0.022	0.075	0.265	0.812	2.131	5.621	13.109	22.586	65.943
$\nu'(q)$	-48.777	-15.717	-8.711	-3.493	-1.096	-0.197	2.632	9.333	120.315
$f(\theta)$	1.118	1.118	1.118	3.354	9.874	15.154	22.444	28.657	50.489

Table 10: Percentiles of t-Statistics across Quantities and Villages for Beans (Linear Specification)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.005	0.053	0.826	3.018	8.307	18.076	41.567	55.740	86.208
$\gamma(\theta(q))$	2.838	7.235	13.672	41.614	157.055	561.841	2144.465	5528.561	7.0×10^4
$\theta(q)$	0.000	0.153	0.338	0.980	2.584	5.454	10.271	13.884	21.219
$\nu'(q)$	-17.063	-9.213	-6.082	-3.085	-1.277	-0.207	3.197	10.292	41.422
$f(\theta)$	1.118	1.118	1.581	5.244	11.854	19.007	25.922	30.166	36.742

Table 11: Percentiles of t-Statistics across Quantities and Villages for Sugar (Linear Specification)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.036	0.282	2.382	5.921	17.718	42.050	68.605	105.833	147.383
$\gamma(\theta(q))$	1.223	4.624	9.748	36.081	134.470	557.324	3948.123	2.1×10^4	5.9×10^5
$\theta(q)$	0.015	0.155	0.345	1.322	4.044	8.708	16.235	23.071	62.843
$\nu'(q)$	-38.690	-15.960	-10.907	-5.386	-1.955	-0.416	0.744	7.616	45.952
$f(\theta)$	1.118	1.118	1.581	6.661	12.966	18.941	26.292	31.682	63.498

Table 12: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Rice (Linear)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	0.610	3.486	6.944	14.628	51.198	142.166	1911.203	7074.768	1.8×10^5
	p_{50}	1.433	11.990	17.459	47.653	244.704	683.711	1.2×10^4	7.3×10^4	3.7×10^5
	p_{75}	17.025	42.254	65.914	333.436	1287.829	5931.734	9.4×10^4	1.0×10^6	1.3×10^8
$\theta(q)$	p_{25}	0.007	0.058	0.137	0.480	1.321	3.313	7.843	12.961	27.037
	p_{50}	0.008	0.080	0.269	0.975	1.856	4.823	10.258	15.960	43.551
	p_{75}	0.118	0.303	0.735	1.567	3.209	7.611	19.342	34.844	76.220
$\nu'(q)$	p_{25}	-72.804	-20.151	-13.900	-5.614	-1.888	-0.883	-0.273	-0.117	-0.025
	p_{50}	-39.347	-14.474	-6.112	-2.895	-1.067	-0.295	-0.024	0.564	111.678
	p_{75}	-20.170	-8.431	-5.135	-1.763	-0.339	1.235	6.809	16.707	268.915
$f(\theta)$	p_{25}	1.118	1.118	1.118	1.936	3.708	7.415	13.387	26.830	31.145
	p_{50}	5.000	5.809	6.778	8.588	12.196	16.657	27.177	28.657	50.236
	p_{75}	5.590	7.104	8.062	9.638	13.145	17.248	27.178	29.118	50.506

Table 13: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Beans (Linear)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	2.367	6.248	8.921	16.583	42.546	136.932	190.165	296.813	937.786
	p_{50}	13.283	28.548	39.386	58.733	119.021	278.601	422.660	615.719	9411.721
	p_{75}	48.148	79.329	136.522	246.220	526.747	1038.497	2927.352	5545.967	9.2×10^4
$\theta(q)$	p_{25}	0.000	0.135	0.254	0.642	1.596	3.182	6.343	12.835	16.059
	p_{50}	0.000	0.158	0.402	1.072	2.744	5.465	8.930	13.431	18.247
	p_{75}	0.000	0.443	0.894	1.813	3.948	6.954	11.479	17.412	21.527
$\nu'(q)$	p_{25}	-17.318	-9.501	-7.744	-4.866	-2.612	-1.241	-0.614	-0.422	-0.010
	p_{50}	-10.875	-6.196	-4.865	-2.601	-1.313	-0.365	-0.010	0.419	9.995
	p_{75}	-10.453	-4.424	-2.919	-1.603	-0.464	1.225	5.346	11.147	39.453
$f(\theta)$	p_{25}	1.118	1.581	1.809	3.162	5.208	8.761	12.247	16.508	33.015
	p_{50}	3.859	7.416	8.441	10.548	15.012	20.520	27.613	33.015	37.450
	p_{75}	7.583	9.014	10.061	12.748	17.212	21.689	27.613	33.015	37.450

Table 14: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Sugar (Linear)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	0.770	2.208	6.424	15.417	40.989	148.159	415.754	977.245	1.4×10^5
	p_{50}	1.492	12.801	25.517	51.241	101.134	305.741	1783.323	5404.830	1.3×10^6
	p_{75}	5.522	49.506	80.162	138.920	335.900	1076.368	8781.182	1.9×10^4	1.5×10^7
$\theta(q)$	p_{25}	0.015	0.090	0.231	0.710	2.441	6.748	10.523	11.369	21.648
	p_{50}	0.015	0.190	0.429	1.573	4.178	7.736	14.294	18.344	28.485
	p_{75}	0.018	0.333	0.729	2.533	6.565	10.463	20.306	28.347	62.843
$\nu'(q)$	p_{25}	-38.690	-21.273	-13.930	-8.272	-4.572	-1.676	-0.573	-0.198	-0.018
	p_{50}	-20.469	-10.970	-8.153	-4.534	-2.247	-0.953	-0.242	-0.084	0.488
	p_{75}	-11.598	-6.389	-4.141	-2.189	-0.759	-0.076	2.067	11.057	62.436
$f(\theta)$	p_{25}	1.118	1.118	1.118	4.127	7.349	10.782	15.969	24.564	63.433
	p_{50}	5.509	6.519	7.983	11.883	15.906	21.125	26.972	36.573	63.498
	p_{75}	6.124	8.367	9.287	13.178	17.401	24.005	27.605	37.006	63.498

B.3.2 Standard Model

In Figures 15-17, we plot the estimates of marginal cost, $c'(Q)$, the distribution of types, as captured by its support, $\theta(q)$, and its reverse hazard rate, $f(\theta)/F(\theta)$, and the base marginal utility function, $\nu'(q)$, when we set $\gamma(\theta(q))$ equal to one for all quantities in all villages, as in the standard model of Maskin and Riley (1984). For brevity, we use the abbreviation "SM" to refer to the standard model. As mentioned, the range of the support of consumer types is less wide, and base marginal utility is less convex, according to the standard model than according to our model.

Figure 15: Estimated Marginal Cost (SM)

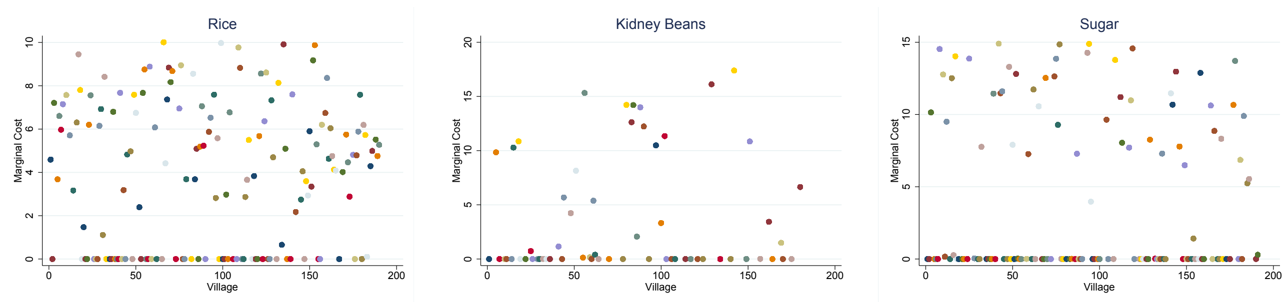


Figure 16: Estimated Distribution of Types (SM)

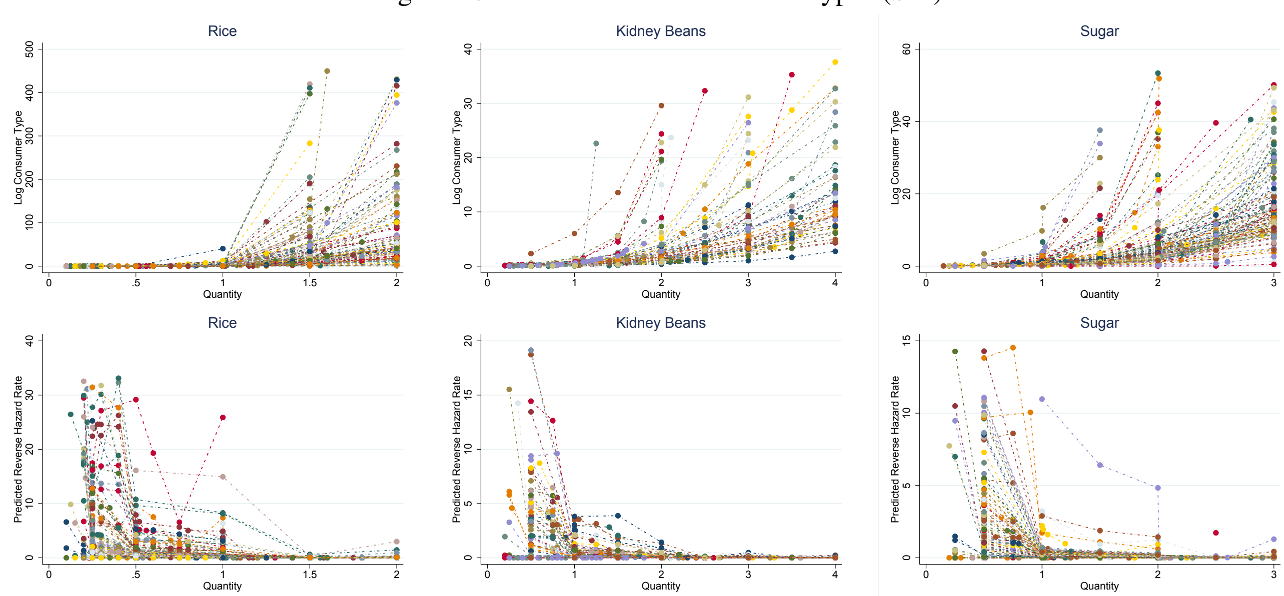


Figure 17: Estimated Base Marginal Utility (SM)

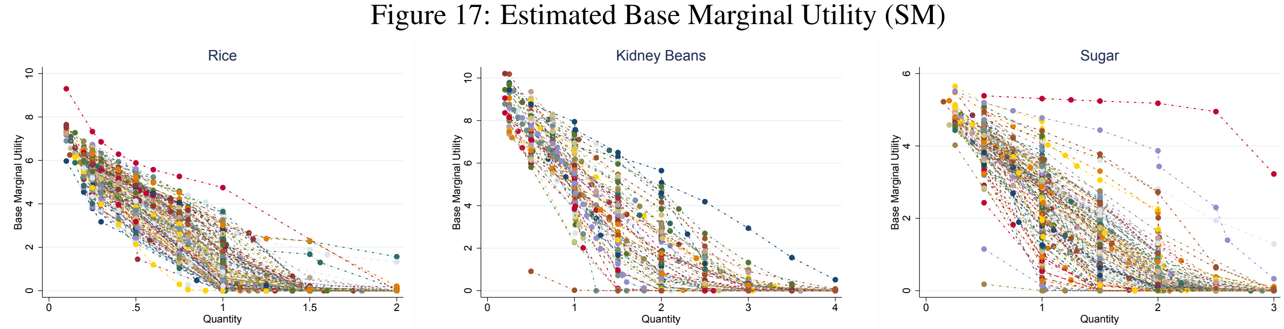


Table 15: Percentiles of t-Statistics across Quantities and Villages for Rice (SM)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.029	0.465	1.569	5.737	13.606	24.958	46.042	57.108	123.646
$\gamma(\theta(q))$	1.177	3.891	6.062	22.413	112.474	1652.991	3.8×10^4	3.6×10^5	7.1×10^7
$\theta(q)$	0.045	0.113	0.363	0.967	2.735	7.174	22.840	46.671	138.463
$\nu'(q)$	-109.942	-26.071	-13.560	-3.979	-0.980	-0.008	10.468	32.654	189.972
$f(\theta)$	1.118	1.118	1.118	3.665	9.671	15.163	24.950	29.262	50.702

Table 16: Percentiles of t-Statistics across Quantities and Villages for Beans (SM)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.015	0.047	0.165	2.363	7.364	15.099	26.034	35.793	60.875
$\gamma(\theta(q))$	2.585	5.642	10.147	31.358	156.134	492.064	3603.320	2.2×10^4	1.6×10^5
$\theta(q)$	0.000	0.135	0.519	1.458	3.513	7.078	11.241	15.223	23.309
$\nu'(q)$	-15.049	-9.306	-7.152	-3.434	-1.202	0.615	13.109	30.124	80.267
$f(\theta)$	1.118	1.118	1.581	5.244	11.565	18.473	24.239	27.613	36.652

Table 17: Percentiles of t-Statistics across Quantities and Villages for Sugar (SM)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.004	0.012	0.081	3.668	11.742	30.069	72.451	129.396	274.517
$\gamma(\theta(q))$	1.166	4.322	8.200	32.756	170.466	575.413	3649.744	1.8×10^4	1.7×10^6
$\theta(q)$	0.002	0.225	0.414	1.127	2.683	5.862	12.785	18.854	49.134
$\nu'(q)$	-28.371	-12.893	-8.537	-4.123	-1.345	-0.326	2.810	9.445	51.105
$f(\theta)$	1.118	1.118	1.118	6.800	12.942	18.874	24.520	27.407	35.777

Table 18: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Rice (SM)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	0.940	2.901	4.969	8.183	23.735	114.980	3113.801	2.9×10^4	6.3×10^5
	p_{50}	1.502	10.257	15.798	32.789	84.607	635.232	9581.298	1.9×10^5	1.8×10^6
	p_{75}	8.125	24.850	44.102	125.823	397.046	3299.666	7.4×10^4	2.4×10^6	2.2×10^8
$\theta(q)$	p_{25}	0.024	0.087	0.281	0.704	1.725	4.215	14.759	38.274	66.776
	p_{50}	0.082	0.155	0.495	1.171	2.541	6.509	18.461	58.424	99.308
	p_{75}	0.148	0.543	0.833	1.805	4.076	9.827	28.787	77.431	119.838
$\nu'(q)$	p_{25}	−109.942	−49.192	−22.981	−6.238	−2.203	−0.848	−0.243	−0.125	1.199
	p_{50}	−83.125	−33.848	−13.560	−3.826	−0.928	−0.101	1.537	4.799	63.914
	p_{75}	−28.180	−16.361	−5.874	−1.777	−0.092	3.762	24.882	65.786	188.850
$f(\theta)$	p_{25}	1.118	1.118	1.118	2.168	4.330	9.093	15.330	28.657	50.279
	p_{50}	2.739	5.000	6.120	8.281	11.307	15.163	27.161	31.031	50.727
	p_{75}	5.123	6.703	7.583	9.269	12.500	17.746	27.166	31.055	50.727

Table 19: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Beans (SM)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	2.222	4.049	8.732	12.704	34.369	112.622	231.876	295.473	908.156
	p_{50}	4.337	12.143	21.240	41.340	107.275	231.221	436.778	626.987	3431.720
	p_{75}	20.332	65.232	115.150	172.986	411.486	1299.782	2837.624	6372.247	4.9×10^4
$\theta(q)$	p_{25}	0.001	0.120	0.384	0.927	2.628	5.160	8.517	10.301	15.412
	p_{50}	0.024	0.271	0.576	1.458	3.547	6.453	9.597	13.718	21.626
	p_{75}	0.024	0.438	0.907	2.165	4.421	8.372	13.787	17.482	24.552
$\nu'(q)$	p_{25}	−15.421	−12.178	−9.057	−5.783	−2.612	−1.101	−0.306	−0.047	2.284
	p_{50}	−10.922	−7.841	−5.372	−2.967	−1.324	−0.185	1.481	2.389	35.157
	p_{75}	−7.341	−4.222	−3.387	−1.170	0.429	5.723	11.765	21.521	66.214
$f(\theta)$	p_{25}	1.118	1.581	1.809	3.162	5.477	9.417	13.509	16.508	31.605
	p_{50}	4.183	6.349	7.804	10.094	15.026	20.520	27.060	32.365	36.646
	p_{75}	7.045	8.657	9.421	12.298	16.771	21.700	27.458	32.365	36.663

Table 20: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Sugar (SM)

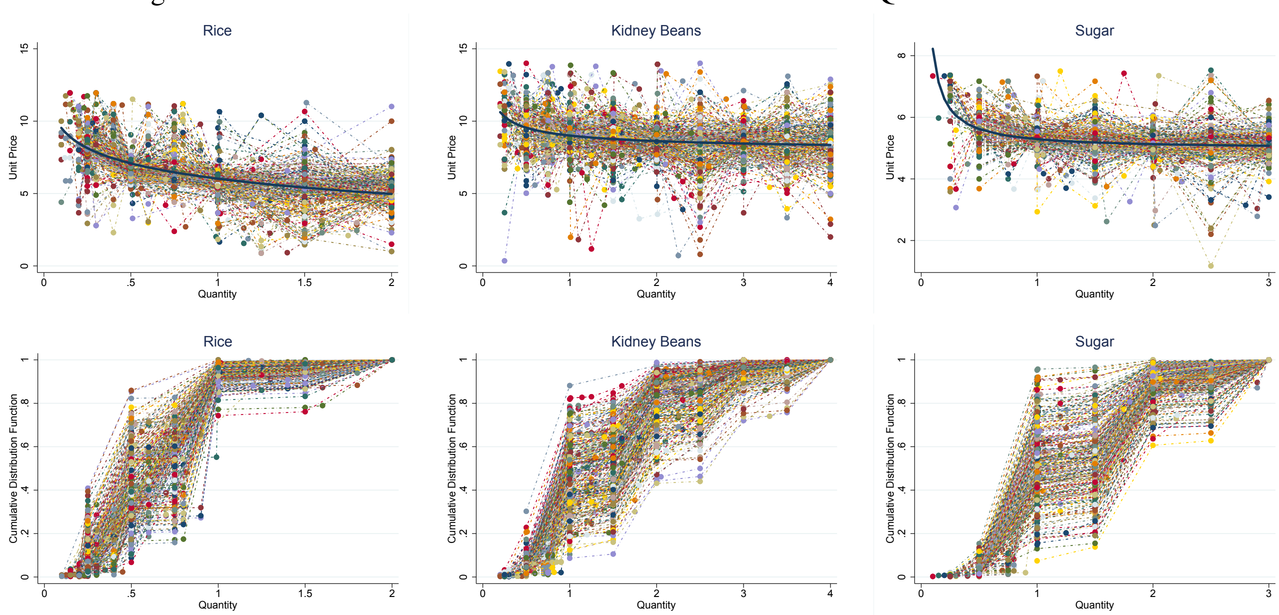
		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	1.248	2.004	5.879	12.244	44.027	191.167	424.775	2054.426	6.9×10^5
	p_{50}	3.344	12.293	20.236	38.504	119.964	289.418	888.764	6884.952	7.3×10^5
	p_{75}	14.400	31.077	39.699	124.092	345.134	886.617	3358.916	6.0×10^4	2.9×10^6
$\theta(q)$	p_{25}	0.000	0.144	0.281	0.906	2.001	3.989	8.619	10.142	22.735
	p_{50}	0.082	0.303	0.797	1.109	2.721	6.396	11.350	16.743	31.355
	p_{75}	0.111	0.402	0.978	1.869	3.790	8.155	17.517	25.216	113.818
$\nu'(q)$	p_{25}	−95.314	−16.733	−12.057	−6.204	−2.784	−1.144	−0.564	−0.326	−0.094
	p_{50}	−21.357	−9.803	−7.400	−3.830	−1.654	−0.730	−0.208	0.014	0.810
	p_{75}	−11.786	−5.866	−4.410	−1.864	−0.635	0.474	3.811	8.184	70.169
$f(\theta)$	p_{25}	1.118	1.118	1.350	3.873	6.982	10.062	13.229	19.170	23.372
	p_{50}	3.741	7.746	9.152	12.349	15.652	20.767	25.125	26.972	34.424
	p_{75}	6.801	8.056	9.485	13.038	17.045	22.220	27.407	32.215	36.552

B.3.3 Estimates Based on Villages Defined as Mexican Localities

So far we have mostly focused on data and estimates from villages defined at the level of the Mexican *municipality*. For a comparison, here we report statistics from the Progresa sample of Mexican localities as well as estimates of the model's primitives and the cumulative multiplier obtained when villages are defined at the level of the Mexican *locality*. Note that the unit price schedules and the empirical cumulative distribution function of quantities for each good are very similar across villages defined as localities and defined as municipalities, as apparent by comparing Figure 18 to Figure 1. We successfully estimate the model's primitives on 326, 333, and 316 of the villages defined as localities for the linear specification of the index of the multiplier function and on 341, 342, and 351 of these villages for the quadratic specification of the index of the multiplier function for rice, kidney beans, and sugar, respectively.

As apparent from Figures 19-23, compared to the analogous ones for villages defined as municipalities, estimation results are similar across the two definitions of villages except for three main differences. First, marginal cost estimates obtained from villages defined as localities are much less dispersed across villages than those obtained from villages defined as municipalities. Second, the support of consumer types is less wide when villages are defined as localities than when they are defined as municipalities. Third, the base marginal utility function for sugar is flatter for some villages, when villages are defined as localities than when they are defined as municipalities. Also, estimates do not differ much across the linear and quadratic specifications of the multiplier function when villages are defined as localities, as is the case for villages defined as municipalities. Counterfactual results about the impact of nonlinear pricing relative to first-best and linear pricing based on estimates from villages defined as localities are very similar to those from villages defined as municipalities, and are therefore omitted for brevity. See the Supplementary Appendix for the estimates of the probability density function of consumer types and for the t-statistics of the model's estimates.

Figure 18: Unit Prices and Cumulative Distribution Function of Quantities from Localities



C The Effect of Income Transfers

In Table B.3, we report the results of an exercise analogous to that discussed in Section 5.4, except that we now allow for fixed effects in all regressions to capture the variability of marginal cost across localities—we also repeat Table 2 in the top panel for ease of comparison. As noted in Section 5.4, results obtained when fixed effects are included are quite similar to those obtained when fixed effects are omitted. This finding is consistent with the relatively small range of variation of estimated marginal cost across localities for each good; see Figure 19.

Figure 19: Estimated Marginal Cost (Linear and Quadratic)

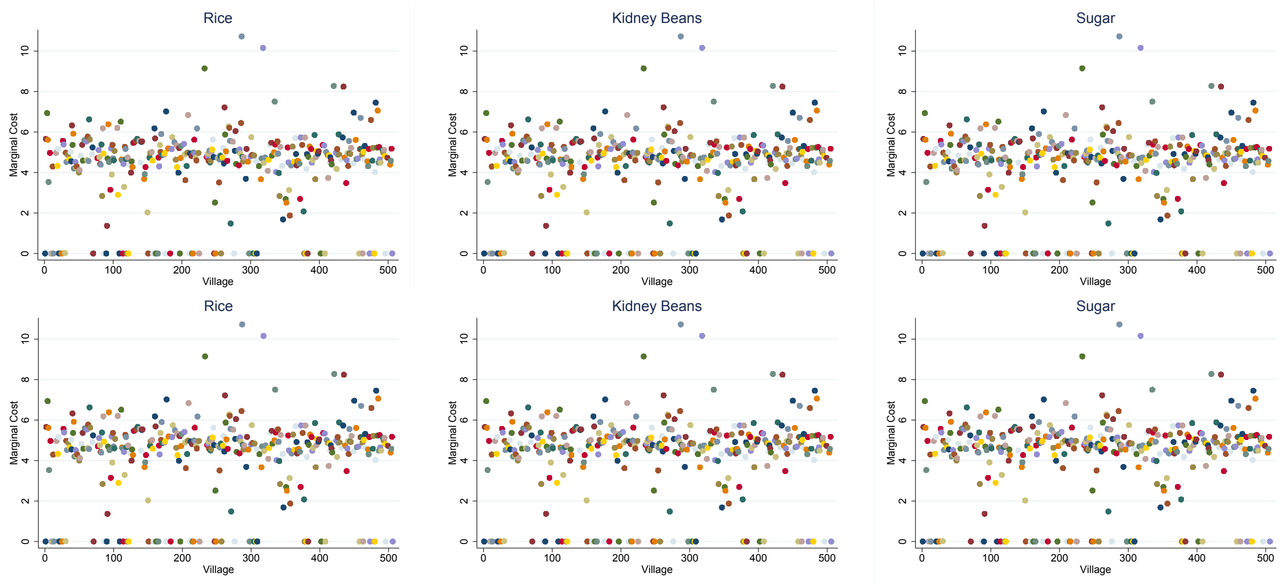


Figure 20: Estimated Multiplier on Participation (or Budget) Constraints from Localities (Linear and Quadratic)

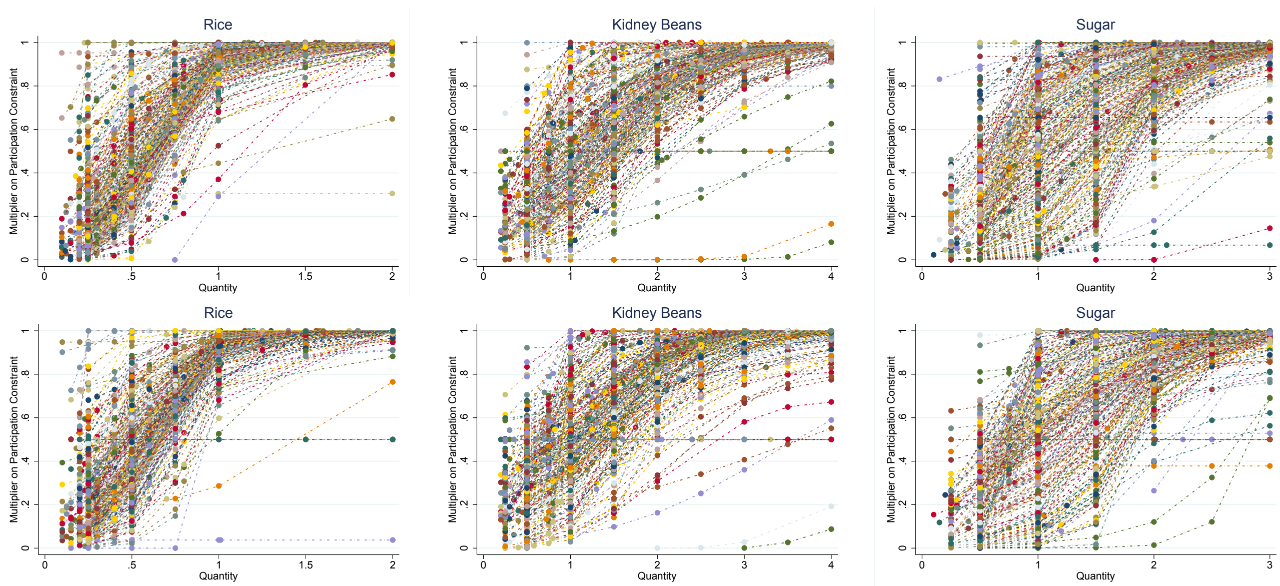


Figure 21: Estimated Distribution of Types from Localities (Linear)

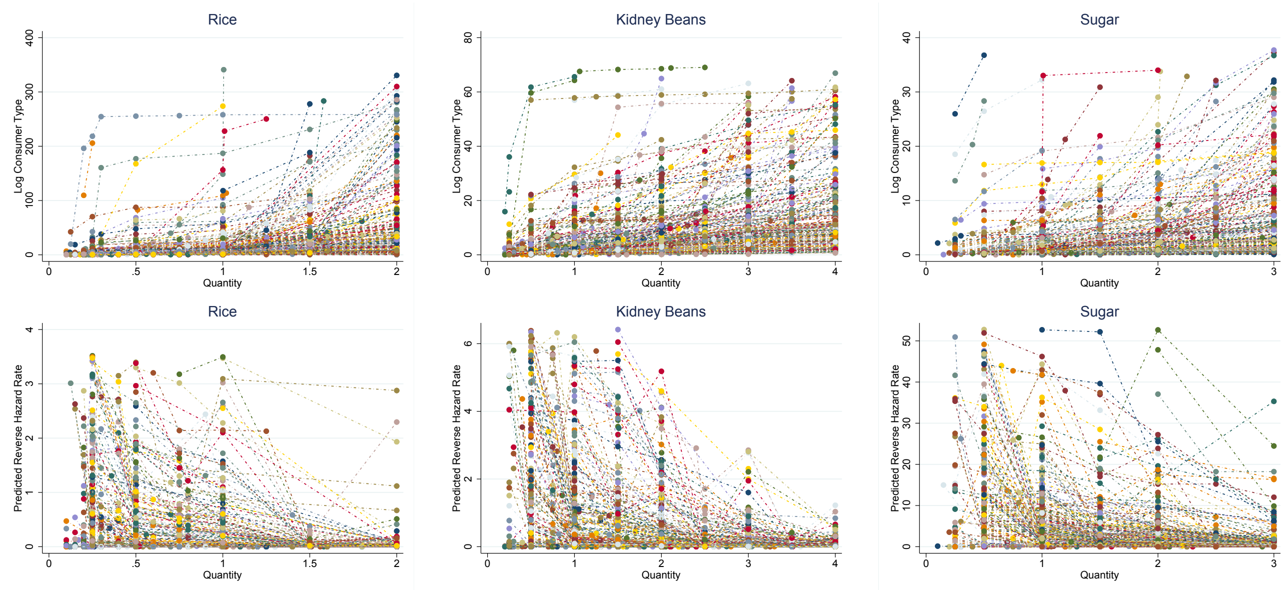


Figure 22: Estimated Distribution of Types from Localities (Quadratic)

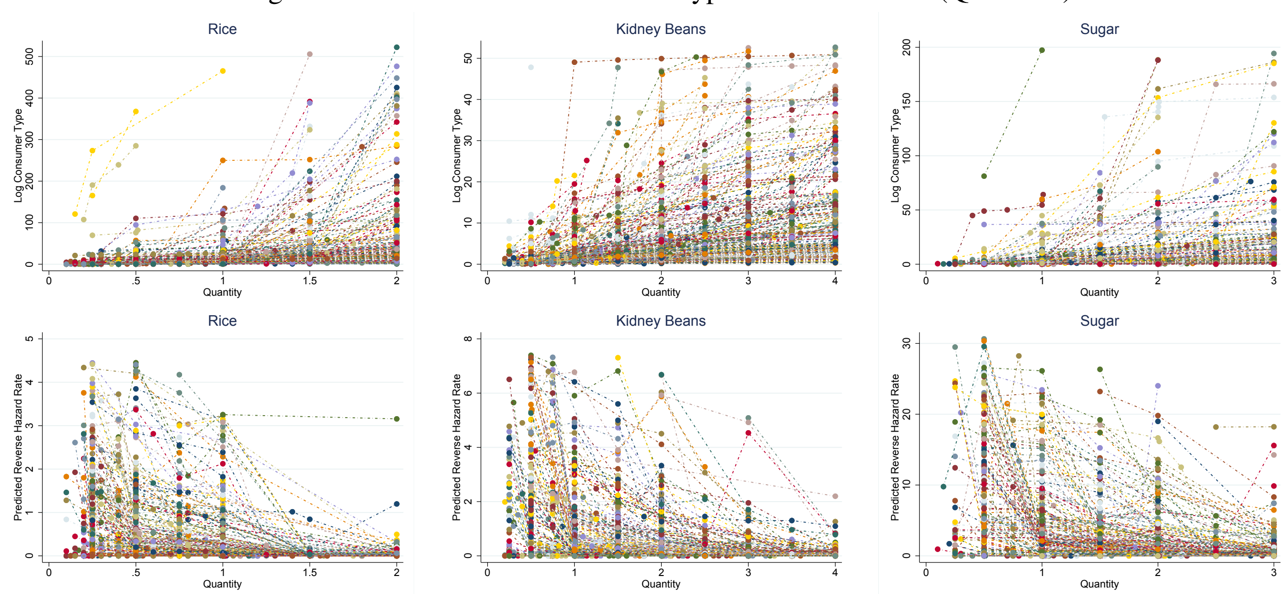


Figure 23: Estimated Base Marginal Utility from Localities (Linear and Quadratic)

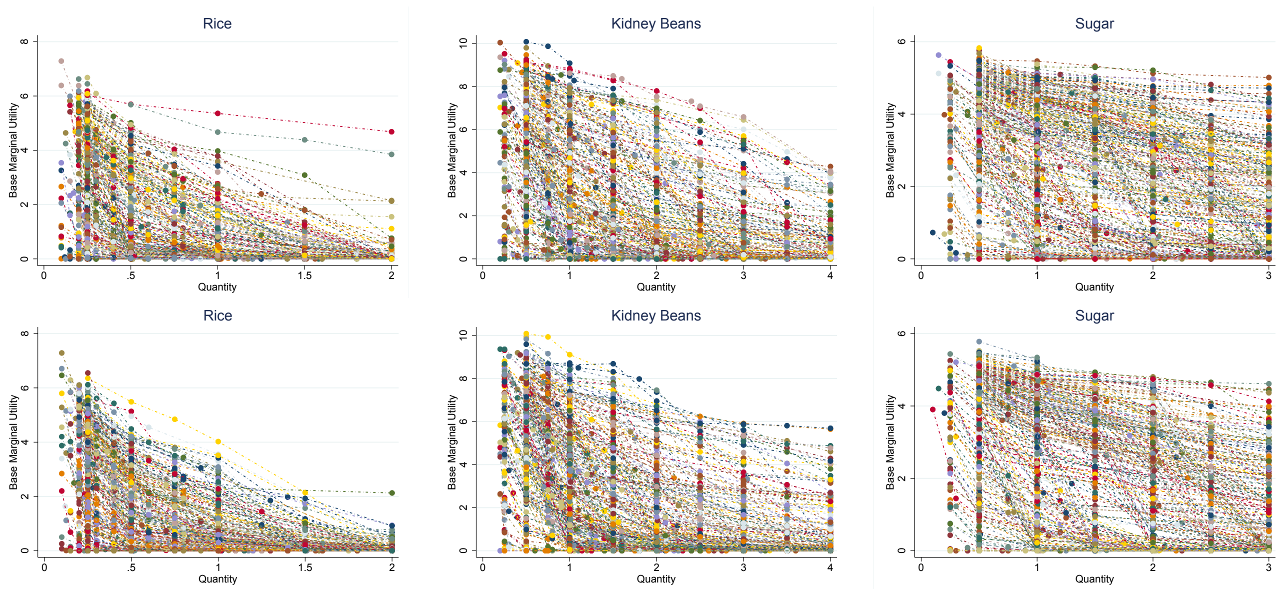


Table B.3: Effect of Cash Transfers on Prices

	Rice Unit Values				Kidney Beans Unit Values				Sugar Unit Values			
	1	2	3	4	1	2	3	4	1	2	3	4
Intercept	1.947***	1.868***	1.864***	1.869***	2.375***	2.420***	2.421***	2.423***	1.743***	1.775***	1.777***	1.769***
	(0.0105)	(0.0127)	(0.00711)	(0.0121)	(0.00940)	(0.00952)	(0.00767)	(0.00998)	(0.00403)	(0.00660)	(0.00448)	(0.00618)
Treatment	−0.00750	−0.0123		−0.00663	−0.0142	0.00600		−0.00494	0.00556	0.0265**		0.0120
	(0.0112)	(0.0149)		(0.0142)	(0.0117)	(0.0127)		(0.0112)	(0.00485)	(0.00909)		(0.00660)
$\log(q)$		−0.212***	−0.181***	−0.161***		−0.114***	−0.172***	−0.160***		−0.123***	−0.182***	−0.168***
		(0.0175)	(0.0324)	(0.0413)		(0.0109)	(0.0281)	(0.0294)		(0.0183)	(0.0172)	(0.0220)
$\frac{1-G(q)}{g(q)}$			−0.0110**	−0.0111**			−0.00505***	−0.00503***			−0.00275*	−0.00273*
			(0.00423)	(0.00422)			(0.00110)	(0.00110)			(0.00107)	(0.00107)
$\log(q) \times \frac{1-G(q)}{g(q)}$			−0.00951*	−0.00943*			−0.00555*	−0.00554*			−0.00546*	−0.00543*
			(0.00469)	(0.00469)			(0.00219)	(0.00219)			(0.00219)	(0.00218)
$\log(q) \times$ Treatment		−0.0441		−0.0270		−0.0399*		−0.0152		−0.0617*		−0.0173
		(0.0255)		(0.0228)		(0.0191)		(0.0136)		(0.0271)		(0.0188)
$\log(q)^2$			0.0458	0.0457			0.0526*	0.0523*			0.0813***	0.0810***
			(0.0398)	(0.0395)			(0.0251)	(0.0250)			(0.0174)	(0.0173)
$\left[\frac{1-G(q)}{g(q)}\right]^2$			0.0000243	0.0000261			0.00000484	0.00000473			−0.0000173	−0.0000171
			(0.0000469)	(0.0000470)			(0.0000213)	(0.0000212)			(0.0000242)	(0.0000242)
1999-March	0.0500***	0.0356***	0.0355***	0.0362***	−0.0827***	−0.0872***	−0.0867***	−0.0865***	0.00888	0.0118*	0.0115*	0.0117*
	(0.00858)	(0.00757)	(0.00788)	(0.00777)	(0.00785)	(0.00743)	(0.00744)	(0.00743)	(0.00551)	(0.00528)	(0.00518)	(0.00516)
1999-Nov	0.0501***	0.0355***	0.0357***	0.0362***	−0.115***	−0.114***	−0.111***	−0.111***	0.0549***	0.0583***	0.0592***	0.0594***
	(0.00772)	(0.00744)	(0.00749)	(0.00739)	(0.00734)	(0.00722)	(0.00719)	(0.00720)	(0.00463)	(0.00439)	(0.00429)	(0.00427)
2000-Nov	0.0302***	0.0213**	0.0229***	0.0217**	−0.196***	−0.203***	−0.205***	−0.201***	0.0307***	0.0283***	0.0330***	0.0303***
	(0.00793)	(0.00679)	(0.00662)	(0.00767)	(0.0122)	(0.0115)	(0.0109)	(0.0114)	(0.00581)	(0.00567)	(0.00518)	(0.00565)
2003	−0.0146	0.00932	0.00536	0.00504	−0.0207	−0.0316**	−0.0348***	−0.0307**	0.170***	0.174***	0.172***	0.169***
	(0.0101)	(0.00996)	(0.00924)	(0.00954)	(0.0110)	(0.0101)	(0.00894)	(0.0101)	(0.00724)	(0.00693)	(0.00611)	(0.00654)
Fixed Effects	N	N	N	N	N	N	N	N	N	N	N	N
R^2 adj.	0.00608	0.140	0.165	0.166	0.0442	0.0946	0.130	0.130	0.0468	0.118	0.201	0.201
Observations	62,368	62,368	62,368	62,368	82,024	82,024	82,024	82,024	91,782	91,782	91,782	91,782

	Rice Unit Values				Kidney Beans Unit Values				Sugar Unit Values			
	1	2	3	4	1	2	3	4	1	2	3	4
Intercept	1.949***	1.873***	1.856***	1.875***	2.378***	2.422***	2.422***	2.424***	1.747***	1.778***	1.760***	1.754***
	(0.0123)	(0.0140)	(0.00754)	(0.0132)	(0.0117)	(0.0114)	(0.00769)	(0.0113)	(0.00591)	(0.00819)	(0.00403)	(0.00685)
Treatment	−0.0136	−0.0342		−0.0272	−0.0101	0.00653		−0.00467	−0.000814	0.0221		0.00795
	(0.0153)	(0.0176)		(0.0165)	(0.0157)	(0.0161)		(0.0146)	(0.00848)	(0.0122)		(0.00972)
$\log(q)$		−0.221***	−0.203***	−0.176***		−0.109***	−0.167***	−0.156***		−0.121***	−0.184***	−0.171***
		(0.0189)	(0.0322)	(0.0418)		(0.0107)	(0.0282)	(0.0298)		(0.0191)	(0.00906)	(0.0175)
$\frac{1-G(q)}{g(q)}$			−0.0128**	−0.0129**			−0.00465***	−0.00464***			0.00606***	0.00604***
			(0.00435)	(0.00433)			(0.00111)	(0.00110)			(0.00136)	(0.00137)
$\log(q) \times \frac{1-G(q)}{g(q)}$			−0.00947*	−0.00938*			−0.00533*	−0.00532*			−0.00687***	−0.00683***
			(0.00464)	(0.00465)			(0.00220)	(0.00219)			(0.00147)	(0.00148)
$\log(q) \times$ Treatment		−0.0512		−0.0354		−0.0418*		−0.0148		−0.0670*		−0.0168
		(0.0262)		(0.0233)		(0.0192)		(0.0137)		(0.0284)		(0.0189)
$\log(q)^2$			0.0435	0.0433			0.0547*	0.0543*			0.120***	0.120***
			(0.0393)	(0.0390)			(0.0253)	(0.0252)			(0.00935)	(0.00934)
$\left[\frac{1-G(q)}{g(q)}\right]^2$			0.0000476	0.0000492			0.00000545	0.00000550			0.00000451	0.00000443
			(0.0000477)	(0.0000480)			(0.0000216)	(0.0000215)			(0.0000111)	(0.0000110)
1999-March	0.0515***	0.0400***	0.0390***	0.0400***	−0.0882***	−0.0920***	−0.0915***	−0.0912***	0.0105	0.0136**	0.0131*	0.0132*
	(0.00847)	(0.00750)	(0.00783)	(0.00771)	(0.00779)	(0.00742)	(0.00742)	(0.00741)	(0.00543)	(0.00518)	(0.00516)	(0.00516)
1999-Nov	0.0539***	0.0416***	0.0414***	0.0420***	−0.119***	−0.117***	−0.114***	−0.114***	0.0545***	0.0592***	0.0585***	0.0586***
	(0.00763)	(0.00737)	(0.00754)	(0.00742)	(0.00735)	(0.00721)	(0.00722)	(0.00721)	(0.00462)	(0.00448)	(0.00444)	(0.00444)
2000-Nov	0.0336***	0.0310***	0.0249***	0.0305***	−0.204***	−0.208***	−0.210***	−0.206***	0.0342***	0.0321***	0.0338***	0.0325***
	(0.00925)	(0.00840)	(0.00688)	(0.00849)	(0.0129)	(0.0123)	(0.0110)	(0.0123)	(0.00663)	(0.00650)	(0.00526)	(0.00650)
2003	−0.00921	0.0219*	0.0112	0.0177	−0.0298*	−0.0379***	−0.0414***	−0.0374***	0.173***	0.177***	0.171***	0.169***
	(0.0102)	(0.0101)	(0.00921)	(0.00981)	(0.0117)	(0.0110)	(0.00900)	(0.0110)	(0.00772)	(0.00747)	(0.00614)	(0.00708)
Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
R^2 adj.	0.0061	0.1394	0.1651	0.1651	0.0442	0.0946	0.1298	0.1301	0.0468	0.1179	0.2012	0.2013
Observations	62,368	62,368	62,368	62,368	82,024	82,024	82,024	82,024	91,782	91,782	91,782	91,782

Note: the superscript * stands for $p < 0.05$, the superscript ** for $p < 0.01$, and the superscript *** for $p < 0.001$. Standard errors are clustered at the locality level.