Nonlinear Pricing in Village Economies*

Orazio Attanasio† Elena Pastorino‡

June 2019

Abstract

This paper examines the price of basic staples in rural Mexico. We document that nonlinear pricing in the form of quantity discounts is common, that quantity discounts are sizable, and that the well-known conditional cash transfer program Progresa has significantly increased quantity discounts, although the program, as documented in previous studies, has not affected unit prices on average. To account for these patterns, 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, because of 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 rely on known results (Jullien (2000)) to characterize the equilibrium price schedule, which is nonlinear in quantity. We analyze the effect of nonlinear pricing on market participation as well as the impact of a market-wide transfer, analogous to the Progresa one, when consumers are differentially constrained. We show that the model is structurally identified from data on prices and quantities from a single market under common assumptions. We estimate the model using data from municipalities and localities in Mexico on three commonly consumed commodities. Interestingly, we find that nonlinear pricing is beneficial to a large number of households, including those consuming small quantities, relative to linear pricing mostly because of the higher degree of market participation that nonlinear pricing induces. We also show that the Progresa transfer has affected the slope of the price schedule of each commodity, which has become steeper as consistent with our model, leading to an increase in the intensity of price discrimination. Finally, we show that a reduced form of our model, in which the size of quantity discounts depends on the hazard rate of the distribution of quantities purchased, explains a large fraction of the shift in price schedules induced by the program.

Keywords: Nonlinear pricing; Budget Constraints; Cash Transfers; Structural Estimation
JEL Codes: D42, D43, D82, I38, O12, O13, O22

*We are grateful to Aureo de Paola, Bruno Jullien, Toru Kitagawa, Aprajit Mahajan, Costas Meghir, Holger Sieg, and Frank Wolak for helpful suggestions and several seminar audiences for comments. Sam Bailey, Ross Batzer, Javier Brugues-Rodriguez, Weixin Chen, Eugenia Gonzalez Aguado, and Sergio Salgado have provided superlative research assistance. We especially thank Brian Albrecht, Adway De, Kai Ding, Keyvan Eslami, and Rishabh Kirpalani for many useful comments and Joan Gieseke for her invaluable editorial assistance. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

†University College London, Institute for Fiscal Studies, CEPR, FAIR at NHH, and NBER.

‡Stanford University, Hoover Institution, and Federal Reserve Bank of Minneapolis.

1 Introduction

Quantity discounts in the form of unit prices declining with quantity are pervasive in developing countries. McIntosh (2003), for instance, documents differences in the price of drinking water paid by poor and rich households in the Philippines, whereas Pannarunothai and Mills (1997) and Fabricant et al. (1999) report similar differences in the price of health care and services in Sierra Leone and Thailand. Attanasio and Frayne (2006) show evidence that households purchasing basic staples in Colombian villages face price schedules rather than linear prices: within a village, richer households buy larger quantities of the same goods that poorer households buy but pay substantially lower unit prices. This evidence is often interpreted as suggesting that nonlinear pricing has undesirable distributional implications.

This view is consistent with the predictions of the nonlinear pricing model of Maskin and Riley (1984), which we refer to as the *standard model*. This model interprets quantity discounts as arising from a seller's incentive to screen consumers by their marginal willingness to pay for a good through the offer of multiple price and quantity combinations. A key insight of this model 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 the first best but also that underconsumption tends to be more severe for consumers of smaller quantities.

The standard model, however, assumes that consumers differ only in their tastes, are unconstrained in their ability to pay for a good, and have access to similar alternatives to purchasing from a particular seller.[1] 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. As such, the standard model abstracts from crucial features of markets in developing countries, especially those for basic staples, in which households typically spend a large fraction of their incomes, face subsistence constraints on consumption, and have access to several alternative consumption possibilities including self-production and highly subsidized government stores. Any such realistic dimension of heterogeneity across consumers, by affecting consumption, may naturally have important consequences for the welfare implications of any pricing scheme.

Pricing behavior in developing countries has received little attention so far, though. Indeed, when Progresa, one of the first conditional cash transfer programs, was introduced in rural Mexico in 1997, policy makers were concerned about the possibility that a substantial part of the transfers to households associated with the program would be appropriated by shopkeepers in targeted villages through price increases. For this reason, several studies have analyzed the effect of transfers on the *average* unit price of commodities but have consistently found no impact. For instance, Hoddinott et al. (2000) 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) dismiss the possibility that their results are mediated by changes in local unit prices when they assess the impact of Progresa on the consumption of non-eligible households. Although Progresa has not affected unit prices on average, in the presence of nonlinear pricing, however, the program

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

may have resulted in differential changes in the unit prices of different quantities and so may have had large undetected distributional effects.

To analyze the determinants of quantity discounts and evaluate the impact of income transfers in their presence 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 and absolute ability to pay for a good. The model also incorporates a rich set of alternatives to purchasing in a particular market that vary across consumers. We characterize nonlinear pricing in this model and investigate the impact of income transfers on prices and consumption. We estimate the model on data from the Progresa evaluation surveys, which the model fits well, and use it to empirically examine the impact of Progresa on prices. Specifically, we document large quantity discounts for basic staples in rural Mexico. We also find that Progresa has had a significant effect on unit prices by leading to an increase in the magnitude of quantity discounts but that this effect cannot be detected without accounting for the dependence of unit prices on quantity.

The paper makes four contributions. First, we show that when facing subsistence constraints, consumers can be 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* with respect to a seller's good, and their constraints depend on their preferences and incomes. Although this class of models has been considered to be intractable in general, we show that 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 nonlinear pricing when consumers are budget constrained.

Second, we prove that the primitives of the model are identified just from information on the distribution of prices and quantities from one market. The intuition behind this result is simple. According to the model, a seller sets prices to discriminate among consumers with different tastes and budgets. Therefore, a seller's price schedule depends on the distribution of consumers' characteristics. Since the distribution of consumers' characteristics is reflected in the observed distribution of quantities purchased, this latter distribution, together with the price schedule, can be used to recover the determinants of prices and consumption, in particular the distribution of consumers' preferences. The estimation approach we propose relies on a seller's optimality conditions, which imply that the difference between prices and marginal cost depends on the difference between the cumulative multiplier associated with consumers' participation or budget constraints and the cumulative distribution function of consumers' characteristics. Based on this relationship, this approach also allows us to identify consumers whose constraints bind and so distinguish among different versions of our model, including the standard model, which is nested within our model.

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 Progresa, which have been extensively analyzed. The estimates of the model's primitives satisfy the model's restrictions on the inverse relationship between marginal utility and quantity purchased as well as the monotonicity of the reverse hazard rate of the distribution of consumers' unobserved marginal willingness to pay, without being imposed.

Fourth, we study the impact of the Progresa transfer on prices. We document that the unit prices of basic staples in the villages we study are highly nonlinear in quantity in that the unit prices of smaller quantities are higher than the unit prices of larger quantities. Like previous studies, we estimate no effect of the Progresa transfer on average unit prices. However, we find that the transfer has increased the (absolute value of the) slope of unit prices with quantity and so has induced an increase in the intensity of price discrimination, which has affected both household beneficiaries of the program and non-eligible households. Finally, using an approximate reduced-form of the optimality conditions for seller behavior, we show that our model can account for the change in price schedules induced by Progresa that we document.

The intuition behind some of our theoretical results is simple and deserves a mention, in particular the impact of nonlinear pricing on market participation and the impact of income transfers on prices. In terms of participation, we show that nonlinear pricing can be a more efficient mechanism than linear pricing in that it leads naturally to greater market inclusion when consumers are differentially constrained. 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 who purchase smaller quantities, or, more generally, with consumers who have access to especially attractive outside options. Such consumers would be excluded from the market under linear pricing because to induce these consumers to participate, a seller would need 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. Thus, including these consumers typically would not be profitable under linear pricing.

As for the impact of income transfers on prices, our model implies 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, we show that income transfers that are more generous for poorer households, who tend to purchase smaller quantities, lead to greater consumption but can also induce an increase in the degree of price discrimination, thereby exacerbating 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. Interestingly, a large fraction of consumers of small and large quantities consume *above* the first best rather than below it, as often argued and implied by the standard model. Moreover, when comparing observed nonlinear pricing to a counterfactual scenario in which sellers cannot price discriminate, we find that linear pricing leads to smaller consumer surplus and lower consumption for most consumers with low to intermediate valuations for two of the three goods we consider, including consumers of the smallest quantities. In particular, a large fraction of such consumers would be excluded from the market under linear pricing and thus benefit from nonlinear pricing. On the contrary, consumers of large quantities tend to be better off under linear pricing.

Unlike the existing literature, which has examined the impact of transfers on unit prices ignoring their nonlinearity, when we evaluate the impact of Progresa on unit prices, we explicitly account for their variation across quantities and allow the program to affect the entire price schedule. As discussed, our model implies that income transfers to consumers affect unit prices, as sellers adjust their price schedules in response to consumers' higher incomes. In line with the model's implications, we find that the schedule of unit prices has become significantly *steeper* after transfers have been introduced, with lower unit prices for the largest quantities in villages receiving cash transfers. Namely, the transfers have led to an increase in the intensity of price discrimination, thereby reducing some of the benefits from the greater consumption. We also derive a reduced form of the model from consumer and seller optimality conditions that relates unit prices to quantities and the inverse hazard rate of the distribution of quantities purchased in each village, which captures the dispersion of consumers' characteristics. Based on this reduced form, we show that our model can explain the change in unit prices associated with Progresa in that the shift in the price schedules induced by the program, in particular in their slope, arises from a shift in the hazard rate of the distribution of quantities purchased, as predicted by our model.

As for the rest of the paper, in Section 2, we describe our sample of rural villages from the evaluation surveys 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 show that our model is identified and detail our estimation strategy. In Section 5, we discuss our estimates, assess their distributional implications, and evaluate the effect of the Progresa cash transfer on prices. Finally, Section 6 concludes the paper.

2 Quantity Discounts: The Case of Mexico

In this section, we provide a description of our data from the Progresa program, present evidence of quantity discounts, and examine the effect of the program on prices.

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, regular visits to health care centers by young children, and attendance of education sessions on nutrition and health by mothers.

Progresa was aimed at marginalized communities identified according to 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 requirements, eligible households had to have access to certain public services and infrastructure, such as schools and health care centers.

In this first phase of the program targeted to rural communities, the Progresa grant consisted of two components. The first one was meant for families with children younger than 6 and was conditioned on children being brought to health care centers with some regularity. The second component was meant for 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 many households, for instance, households with primary school-aged children whose school attendance is very high. For eligible households, the grant was substantial. On average, transfers amounted to 25% of household income and consumption.

Since the first rollout of the program involved 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, called *municipalities*, to be included in the evaluation sample. Each municipality is composed of several localities, not all of which were included in the evaluation surveys. 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. In our empirical exercises, we use the surveys of October 1998, May and November 1999, November 2000, and 2003. We could not use some waves, such as those collected in 1997 and April 2000, because they do not contain information on household expenditure, which we rely on.

The communities included in the evaluation surveys are small—the average number of households in a locality is just over 50—and remote. Households living in these villages are poor; for them, food accounts for a substantial share of consumption. However, not all households within the targeted villages were eligible. Eligibility for the program 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. On average, about 78% of the households of the villages in the evaluation surveys were considered eligible.[2] The level of poverty of communities in the evaluation surveys, though, exhibits substantial variation not only within but also across villages. This heterogeneity is reflected, for instance, in the variability of the rate of eligible households across villages.

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 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 surveyed, therefore allowing us to estimate the entire distribution of quantities and prices in each village, at least for commodities that are commonly purchased. Third, the data are very rich and exhaustive.

The consumption and expenditure module of the surveys contains crucial information for the purpose of our paper. Each household is interviewed and asked to report not only the quantity consumed of each of 36

[2] 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, at around 52%. This assessment led to a slight modification of eligibility rules. We consider these added families as eligible.

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, those acquired through self-production or received as a gift or payment in kind. The food items recorded include fruits and vegetables, grains and pulses, 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 their quality and are purchased and consumed by most households, as explained below.

Given the information available on expenditure and quantities purchased for each recorded item, it is possible to compute their *unit values*, as measured by the ratios of these variables. From now on, we refer to unit values as *unit 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 the nonlinearity in quantity we consider here. However, they find that average and median unit values well approximate local prices collected from local stores, which are available for some commodities in the locality surveys. They also find that unit values closely match national data sources on prices.

The variability in the rate of households eligible for Progresa across villages mentioned above is naturally related to differences in the distribution of quantities purchased across villages. Such heterogeneity could also account for differences in unit prices 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 consumption and, according to the model we propose, the dependence of prices on quantity within each village. As we show in Section 5.4, changes in the distribution of quantities purchased within a village are key to assessing the ability of our model to account for differences in price schedules across villages and so the impact of the Progresa transfer on unit prices. We now turn to analyzing the impact of the Progresa transfer on unit prices.

Evidence of Quantity Discounts and Price Effects of the Transfer Program. Quantity discounts are 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 purchased is as large as -0.11 in their preferred specification. They estimate even larger quantity discounts for different specifications and other goods such as carrots or beans.

Here we first document the existence of discount patterns in Mexico similar to those observed in Colombia and then examine the impact of Progresa on unit prices. We focus on three goods—rice, kidney beans, and sugar—that conform to the assumptions we maintain in the theoretical model. Specifically, we consider goods that are of homogeneous quality so as to minimize the possibility that price differences reflect any heterogeneity in this unmodeled dimension.[3] The goods we choose are not only widely consumed but also storable so that no observed purchase for a household does not necessarily reflect exclusion from the market

[3]Some studies have argued that the dispersion in prices observed in a market for a same good might reflect differences in quality. Deaton (1989), for instance, argues that this might be the case for rice in Thailand. Here we focus on goods for which the assumption of quality homogeneity does not seem unreasonable in our context in light of conversations with program officials. Also, any quality heterogeneity would likely give rise to upward-sloping price schedules, contrary to what we observe.

Table 1: Price Schedules and the Impact of Cash Transfers on Prices (98% Trimming)

	Rice Unit Values			Kidney Beans Unit Values			Sugar Unit Values		
	1	2	3	1	2	3	1	2	3
Intercept	1.866***	1.994***	1.874***	2.473***	2.399***	2.465***	1.832***	1.768***	1.814***
	(0.005)	(0.008)	(0.007)	(0.007)	(0.010)	(0.010)	(0.004)	(0.004)	(0.006)
Treatment		-0.006	-0.008		-0.007	0.010		0.003	0.025***
		(0.009)	(0.008)		(0.012)	(0.013)		(0.005)	(0.007)
$\log(q)$	-0.320***		-0.290***	-0.188***		-0.161***	-0.198***		-0.157***
	(0.007)		(0.009)	(0.007)		(0.009)	(0.009)		(0.010)
$\log(q) \times$ Treatment			-0.038***			-0.035***			-0.053***
			(0.013)			(0.013)			(0.015)
R^2	0.352	0.136	0.353	0.222	0.146	0.223	0.168	0.045	0.170
Observations	69,543	69,543	69,543	93,375	93,375	93,375	103,930	103,930	103,930

Note: * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$. Clustered standard errors. Wave fixed effects included.

but could simply occur because of the timing of the Progresa interview. Hence, the assumption of full market participation we will formulate in our analysis is not implausible. Indeed, the median fraction of households consuming rice across localities in the week preceding the interview is 59%, whereas the corresponding fraction for kidney beans and sugar is 87%. Virtually all of these households purchase these goods rather than producing them or receiving them as a gift or an in-kind payment: the median fraction of purchasing households across localities is 100% for rice, 94% for kidney beans, and 100% for sugar.

We use data from the Progresa waves of October 1998, May and November 1999, November 2000, and 2003, and focus on villages with at least 50 households purchasing the goods of interest. We exclude observations reported in uncommon units of measurement (different from kilos) and trim the top 2% of the observations on quantities purchased and expenditures, expressed relative to their level in October 1998, to limit the influence of outliers. For the three goods of interest, we examine the relationship between unit prices and quantities purchased in each village.

Columns 1 in Table 1 contain estimates of a regression of log real unit prices on log quantities. The different numbers of observations in each row reflect the different numbers of purchases we observe in our sample. In this exercise as well as in that in columns 2, we include wave fixed effects and cluster standard errors at the village (locality) level. The elasticity of unit prices to quantities we estimate is largest in absolute value for rice, -0.320, but it is also sizable for the other two goods, -0.188 for kidney beans and -0.198 for sugar. For each good, this elasticity is statistically different from zero.

In columns 2, we estimate the effect of Progresa on average unit prices by regressing log real unit prices on a constant and a dummy equal to one for transactions occurring in localities targeted by the program. Consistent with studies that have estimated this impact, such as Hoddinott et al. (2000), we do not find any evidence that the Progresa transfer has affected the average unit prices of these three goods.

We complement this evidence on the effect of the program by examining the possibility that cash transfers have modified these price schedules and affected the magnitude of the quantity discounts that we document in columns 1. In particular, we augment the regressions estimated in columns 1 with a dummy for the program, "Treatment," and an interaction term between this dummy and log quantity to let both the intercept and the slope of price schedules vary with the presence of the program. The results of these augmented regressions, presented in columns 3, show that the program has indeed increased the size of quantity discounts and so the nonlinearity of unit prices for each good. Specifically, the slope of the price schedule for each good has increased in absolute value with the program: from -0.320 to -0.328 for rice, from -0.188 to -0.196 for kidney beans, and from -0.198 to -0.210 for sugar. This effect is significant at the 1% for all goods. (For sugar, we also observe a significant positive effect of the program on the intercept of the price schedule.) Thus, the program has been accompanied by an increase in the intensity of price discrimination, as we discuss in detail in Section 5.4.

Market Structure. The model we present in the next section considers a seller facing a heterogeneous population of consumers. Since the model focuses on the behavior of a single seller, ideally, one would like to consider a relatively isolated market with one seller or a small number of them. Using the localities in our data as such markets would then seem natural. However, such an approach would result in few observed transactions in several instances, given the size of the localities and the number of recorded purchases. Moreover, despite some localities being quite isolated, they all belong to a municipality from an administrative point of view 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 on villages defined as municipalities. However, we estimate our model on village markets defined as both municipalities and localities, and obtain fairly similar results for these two definitions of villages, as discussed in Section 5.

The assumption of one or very few sellers is consistent with our 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 markets and itinerant street markets, might be present. These sellers, however, can be considered as characterized by a very different cost structure, and their possible presence in the markets for the goods we study can be interpreted as a degree of competition that is incorporated into households' outside options in our model. Even at the level of the municipality, the number of grocery stores that might sell the goods we consider is very small: the median number is 1 and the mean is 2 across waves. Hence, the supply of the goods we focus on is highly concentrated in each market, so the restriction to one seller per market in our empirical analysis does not seem too strong.

3 Models of Price Discrimination

As just shown, the unit prices of basic staples in rural Mexico 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 arise when a seller screens consumers by their marginal willingness to pay according to the quantities of a good 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 incorporate richer consumption possibilities as an alternative to purchasing from a particular seller, we build on the model of Jullien (2000), which assumes that 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 a number of settings of interest. For instance, consumers in our data have access to a wide range of consumption opportunities: 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.[4]

An important case for our application is 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.[5] 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 the results in Jullien (2000) to a model with budget-constrained consumers and characterize nonlinear pricing in the presence of these constraints. As consumers typically have preferences for multiple goods, we allow for consumers' substitution across them and let subsistence constraints affect the consumption of any good.

As is common in the nonlinear pricing literature, our framework implicitly excludes the possibility of collusion among consumers, for instance, through resale. Anecdotal evidence from Progresa 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.[6]

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 that the seller knows the distribution of θ but does not observe its value for a given consumer or, alternatively, that the seller observes its value but prices contingent on consumers' characteristics are not enforceable or legally permitted. Thus, a seller must post a single price schedule for all consumers, but this schedule can entail

[4] 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 we develop can account for different degrees of seller market power. For example, 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. See the Supplementary Appendix.

[5] The optimal pricing schedule is known for special cases, when, for instance, utility is linear in consumption (see Che and Gale (2000)) or the budget is identical across consumers (see Thomas (2002)).

[6] Conceptually, such a situation arises in the presence of imperfections in contracting between consumers analogous to those between sellers and consumers usually maintained in models of nonlinear pricing. Specifically, in the presence of enforcement, coordination, or transaction costs such as commuting costs, a coalition of consumers may not be able to achieve higher utility for any member than the utility a member obtains by trading with a price-discriminating seller.

different unit prices for different quantities.[7]

Each consumer decides whether to purchase and, if so, the quantity q to buy. When purchasing from the seller, a consumer of type θ obtains utility $v(\theta, q) - t$, with $v(\cdot, \cdot)$ twice continuously differentiable, $v_\theta(\theta, q) > 0$, $v_q(\theta, q) > 0$, and $v_{qq}(\theta, q) \leq 0$. We assume, as is 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, Q. For simplicity, here we maintain that the cost function $c(\cdot)$ is additively separable across consumers; we relax this assumption in the empirical analysis. We denote by $s(\theta, q) = v(\theta, q) - c(q)$ the *social surplus* from quantity q and assume that $s_q(\theta, \cdot)/v_{\theta q}(\theta, \cdot)$ decreases with q, which ensures that the seller's problem admits a unique solution and that first-order conditions are necessary and sufficient to characterize it. This assumption 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, as in Jullien (2000).

Let $\overline{u}(\theta)$ be a consumer's *reservation utility* when the consumer does not purchase from the seller, which is assumed to be absolutely continuous and, unlike in the standard model, can differ across consumers. 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 normalize the seller's reservation profit to zero. We focus on situations in which all consumers trade, so $q = 0$ is interpreted as the limit when the contracted quantity becomes small. Note that if we allowed for consumer exclusion, then the equilibrium price schedule faced by types who participate would be the same as the one we characterize below.

By the revelation principle, a contract between a seller and consumers 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 $c(Q)$ is shorthand for $\int_{\underline{\theta}}^{\overline{\theta}} c(q(\theta)) f(\theta) d\theta$ when $c(\cdot)$ is additively separable. 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 them. The IC constraint of a consumer of type θ is satisfied if

[7] 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).

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, $u'(\theta) = v_\theta(\theta, q(\theta))$. As $v_{\theta q}(\theta, q) > 0$, an allocation is incentive compatible if, and only if, it is *locally* incentive compatible in that $u'(\theta) = v_\theta(\theta, q(\theta))$ (a.e.), the schedule $q(\theta)$ is weakly increasing, 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 conditions

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

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 $\gamma(\overline{\theta}) = 1$.[8]

Jullien (2000) formulates three important assumptions to characterize an optimal menu, namely, *potential separation* (PS), *homogeneity* (H), and *full participation* (FP). As for (PS), note that for each type θ, the first-order condition in (1) defines the optimal quantity $q(\theta)$ as a function of the primitives of the economy and the cumulative multiplier $\gamma(\theta)$, $q(\theta) = l(\gamma(\theta), \theta)$. The quantity $l(\tilde{\gamma}, \theta)$ that satisfies (1) at θ for the arbitrary cumulative multiplier $\tilde{\gamma} \in [0, 1]$ weakly decreases with $\tilde{\gamma}$. Assumption (PS) states that $l(\tilde{\gamma}, \theta)$ weakly increases with θ for all $\tilde{\gamma} \in [0, 1]$. This assumption guarantees that the seller has an incentive to discriminate across consumers and so effectively ensures that the optimal $q(\theta)$ is weakly increasing.[9] Assumption (H) states that there exists a quantity profile $\{\overline{q}(\theta)\}$ such that $\overline{u}'(\theta) = v_\theta(\theta, \overline{q}(\theta))$ and $\overline{q}(\theta)$ is weakly increasing; that is, the allocation with full participation $\{\overline{u}(\theta), \overline{q}(\theta)\}$ is implementable. This assumption ensures that a consumer's IC constraint can be satisfied when the IR constraint binds. Assumption (FP) states that all types participate. Sufficient conditions for (FP) are (H) and $s(\theta, \overline{q}(\theta)) \geq \overline{u}(\theta)$, which guarantee that the seller has an incentive to trade with all consumers.

Jullien (2000) shows that under these three assumptions, 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 $q(\theta)$ continuous and weakly increasing. The solution to the seller's problem is $q(\theta) = l(\gamma(\theta), \theta)$ with associated price $t(\theta)$ and utility $u(\theta)$, which equals $\overline{u}(\theta)$ for consumer types whose IR constraints bind. Note since $q(\theta)$ is continuous, $\gamma(\theta)$ can have mass points only at

[8] See the Supplementary Appendix for details. 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. It is understood that $q(\theta)$ is evaluated taking the left limit at jump points.

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

$\underline{\theta}$ or $\overline{\theta}$, and thus the IR constraints can bind at isolated points only at $\underline{\theta}$ or $\overline{\theta}$. When $\gamma(\theta) = 1$ for all types so that the IR constraints bind only at $\underline{\theta}$, the model reduces to the standard model, in which the IR constraints simplify to $u(\theta) \geq \overline{u}$ with $\overline{u}$ constant.

Observe that by varying the reservation utility schedule, the model can accommodate different degrees of market power for a seller, ranging from the case of perfect competition to that of monopoly. Specifically, when the reservation utility equals the social surplus under the first best for each type, that is, $\overline{u}(\theta) = v(\theta, q^{FB}(\theta)) - c(q^{FB}(\theta))$, the solution to the seller's problem implies $\gamma(\theta) = F(\theta)$ for all consumers so that consumers purchase first-best quantities from the seller. As the reservation utility is lowered from its maximal value under the first best for each type, profits correspondingly increase, thus allowing the model to capture any degree of imperfect competition (across markets). This feature of the model provides an important dimension of flexibility over the standard model for the measurement exercises in later sections.[10]

3.2 A Model with Heterogeneous Budget Constraints

Suppose now that instead of having access to 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 here 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. We will use this model in the next subsection to examine the impact of income transfers on prices and consumption.

Setup. Suppose that consumers have quasi-linear 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, $Y(w)$.[11] 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. Namely, 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. Clearly, this calorie constraint can be rewritten as $z \geq \underline{z}(\theta, q) \equiv [\underline{C}(\theta) - C^q(\theta, q)]/C^z(\theta)$.[12]

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

[10] With $v_\theta(\cdot, \cdot) > 0$ by assumption, (H) implies that $\overline{u}(\theta)$ is monotone since it requires $\overline{u}'(\theta) = v_\theta(\theta, \overline{q}(\theta))$. This feature of the model is consistent with nearly all households in each village purchasing the goods we consider. Our analysis could be extended to the case in which consumers dislike the seller's good, and can be ranked by their distaste for it, with θ replaced by $-\theta$. (H) also requires that $\overline{q}(\theta)$ be weakly increasing and thus that $\overline{u}(\theta)$ be sufficiently convex, which prevents bunching. Observe that (H) naturally holds in a model of seller competition with vertical differentiation. Under this interpretation of our model, here we characterize the best-response problem of any such oligopolist; see the Supplementary Appendix for details.

[11] We implicitly assume that utility is separable across a seller's goods, which are priced independently. See Stole (2007).

[12] 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.

from the seller, the consumer's problem is

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

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 $z\ge \underline{z}(\theta,q)$, the problem in (3) can be restated as

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

where $I(\theta,q,w)$ is the maximal amount the consumer can spend to purchase q units of the seller's good and meet her subsistence constraint.[13] 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, twice continuously differentiable, and weakly increasing in θ and q. An intuition for why $I(\cdot,q,w)$ may increase with θ, and thus $\underline{z}(\cdot,q)$ may decrease with θ, is that if the same calorie intake can be reached through different combinations of goods, a consumer with a greater taste for the seller's good may require less of other goods to achieve it, for instance, because of a greater ability to metabolize the good. The requirement that $I(\theta,\cdot,w)$ increases with q, and so $\underline{z}(\theta,\cdot)$ decreases with q, is equivalent to $C(\theta,\cdot)$ increasing with q and is natural: the greater the amount of the seller's good consumed, the greater the calorie intake. 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.

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

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

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

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

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

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

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

[13] Quasi-linear 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 usually desirable. The latter, however, are unrealistic in the context of our application.

that induces each consumer to purchase and spend her entire budget for the seller's good, that is,

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

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, condition (BCH) helps to ensure that there exists an implementable menu $\{\bar{t}(\theta), \overline{q}(\theta)\}$ that induces all consumers to participate.[14]

Since income affects consumers' purchase behavior in this model, changes in the distribution of consumers' income arising from, for instance, income transfers typically influence a seller's menu. In the IR model, on the contrary, 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 $\overline{u}(\theta)$ is exogenously assumed to depend on the consumer's income. In this sense, the IR model can be considered a "reduced form" of the BC model. We explore the implications of the BC model for the effect of income transfers below when we assess the impact of Progresa on prices and consumption.

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 to a simple constructive argument, would be to define the budget for the seller's good of a consumer of type θ as $I(\theta, \hat{q}(\theta)) = v(\theta, \hat{q}(\theta)) - \overline{u}_{IR}(\theta)$ for any implementable allocation $\{\hat{u}(\theta), \hat{q}(\theta)\}$ in the BC model, where $\overline{u}_{IR}(\theta)$ denotes the reservation utility for a consumer of type θ in the IR problem. Since $\hat{t}(\theta) = v(\theta, \hat{q}(\theta)) - \hat{u}(\theta)$, it is immediate that the BC constraint is equivalent to the IR constraint of the IR problem in this case. Although this approach is intuitive since it 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 at an implementable allocation. As we now show, for the two problems to admit the same solution, 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 at the optimal menu—as long as consumers have enough income to afford the IR allocation $\{u_{IR}(\theta), q_{IR}(\theta)\}$.

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

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

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

[14]Under (FP), the IR' constraints are effectively redundant. Sufficient conditions for (FP), and so for 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. To see why, note 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. (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. Thus, all types participate.

the cumulative multiplier on the budget constraint expressed as $I(\theta, q(\theta)) \geq v(\theta, q(\theta)) - u(\theta)$ and is defined analogously to $\gamma(\theta)$ with derivative $\phi(\theta)$. The first-order conditions of this problem are

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

for each type, along with the complementary slackness condition

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

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 $\Phi(\overline{\theta}) = 1$. Denote by $\{t_{IR}(\theta), q_{IR}(\theta)\}$ the optimal menu and by $\overline{q}_{IR}(\theta)$ the reservation quantity profile in the IR model. We now establish the desired equivalence.

Proposition 1 (*Equivalence of Problems*)**.** *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 can always induce a consumer to buy by offering a large enough quantity for a given price or by charging a low enough price for a given quantity. The IR constraint, though, 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 $t_{IR}(\theta) \leq v(\theta, q_{IR}(\theta)) - \overline{u}_{IR}(\theta)$, which effectively limits a consumer's expenditure on the seller's good. Hence, in this precise sense, the 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. Then, 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 BC constraints bind, namely, those with $\phi(\theta) > 0$. 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 prices, in the two problems coincide.

A natural question is how stringent the assumptions of Proposition 1 are, in particular the condition $I_q(\theta, \overline{q}_{IR}(\theta)) = v_q(\theta, \overline{q}_{IR}(\theta))$. This condition simply implies that if the seller uses the budget schedule $I(\theta, \overline{q}_{IR}(\theta))$ as a price schedule in the BC model when the BC constraints bind, he can induce consumers to demand the same incentive compatible quantities they demand in the IR model when the IR constraints bind. Hence, consumers with binding constraints demand the same quantities in the two models. This condition can easily be satisfied when the utility of a consumer who spends all of her budget on the seller's good in the BC model is chosen to agree with $\overline{u}_{IR}(\theta)$ so that the quantity $\overline{q}(\theta)$ of the (BCH) assumption coincides with

$\overline{q}_{IR}(\theta)$.[15]

Proposition 1 is important for several reasons. First, it provides 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. Second, this result allows us to examine how subsistence constraints affect prices and consumption as well as to evaluate the effect of policies, such as income transfers, that directly affect consumers' ability to pay and budgets. We do so in the next subsection.

3.3 Properties and Implications of Nonlinear Pricing

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 applying to both cases. We now examine the implications of the augmented model for prices and consumption and for the relative desirability of nonlinear and linear pricing. We then consider the version of the augmented model in which consumers have heterogeneous budget constraints to analyze the impact of policies such as income transfers that affect consumers' ability to pay. We maintain, for simplicity, that $v(\theta, q) = \theta\nu(q)$ and $c'(Q) = c > 0$ and focus on the regular case in which optimal and reservation quantity schedules are increasing with the type.[16]

Prices and Consumption. We start by providing sufficient conditions for quantity discounts to arise. Since $q(\theta)$ is increasing, we can define the inverse function $\theta(q)$ and derive the observed price schedule as a function of quantity, $T(q) = t(\theta(q))$. 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 $p(q)$ declines with q, where $p(q) = T(q)/q$ is the unit price of quantity $q = q(\theta)$.[17] Denote by $A(q) \equiv -\nu''(q)/\nu'(q)$ the coefficient of absolute risk aversion evaluated at q. We can then prove the following result.

Proposition 2 (*Quantity Discounts*)**.** *Assume $\nu''(\cdot) < 0$ and $\frac{\partial}{\partial\theta}\left(\frac{1-F(\theta)}{f(\theta)}\right) \leq 0$. If $\frac{F(\theta)}{\theta f(\theta)} \leq \min\left\{1, \frac{\partial}{\partial\theta}\left(\frac{F(\theta)}{f(\theta)}\right)\right\}$ and $\overline{u}''(\theta) \geq \frac{\nu'(\overline{q}(\theta))}{\theta A(\overline{q}(\theta))}$ for each θ, then $T''(q) \leq 0$ for each $q = q(\theta)$.*

Whereas the first two conditions in the proposition are common—the second one is a sufficient condition for assumption (PS) in the standard model—the remaining two are novel. Consider first the condition on $F(\theta)/[\theta f(\theta)]$. Requiring that $F(\theta)/f(\theta) \leq \theta$ simply bounds the rate of increase of $T(q)$ when the seller offers quantities above the first best, as is the case whenever $\gamma < F(\theta)$. Intuitively, $T'(q)$ is related to

[15] As $\overline{u}'_{IR}(\theta) = v_\theta(\theta, \overline{q}_{IR}(\theta))$ by assumption (H) of the IR model, $I_q(\theta, \overline{q}_{IR}(\theta)) = v_q(\theta, \overline{q}_{IR}(\theta))$ holds when $I_\theta(\theta, \overline{q}_{IR}(\theta)) = 0$.

[16] The optimal quantity schedule is increasing if $s_q(\theta, \cdot)/v_{\theta q}(\theta, \cdot)$ decreases with q, as assumed, $s_q(\cdot, q)/v_{\theta q}(\cdot, q)$ increases with θ, $F(\cdot)/f(\cdot)$ increases with θ, and $[1 - F(\cdot)]/f(\cdot)$ decreases with θ. Note that the assumption $v_\theta(\theta, q) > 0$ implies that $\nu(q) > 0$.

[17] A sufficient condition for $p'(q) \leq 0$ is $T''(q) \leq 0$ with $T(0) \geq 0$. To see why, recall that if $f(x)$ is a concave function, then $f(x) \leq f'(x_2)(x - x_2) + f(x_2)$ or, equivalently, $x_2 f'(x_2) \leq f(x_2) + x f'(x_2) - f(x)$ at any point $(x_2, f(x_2))$. If this inequality holds for any x, then it must also hold for $x = 0$, in which case it becomes $x_2 f'(x_2) \leq f(x_2) - f(0)$ provided that $f'(x_2)$ is bounded. Thus, if $f(0) \geq 0$, then $x_2 f'(x_2) \leq f(x_2)$, and so $f(x_2)/x_2$ declines with quantity.

$F(\theta)$ and $f(\theta)$ by (9) and so to the reverse hazard rate of the distribution of consumer types. Quantity discounts in the form of $T''(q) \leq 0$ require the rate of increase of $T(q)$ to decrease with quantity and thus, by (9), $F(\theta)/f(\theta)$ to increase fast enough with θ. The restriction $\partial\left[F(\theta)/f(\theta)\right]/\partial\theta \geq \left[F(\theta)/\theta f(\theta)\right]$ strengthens the usual condition on the distribution of consumer types for assumption (PS) to hold in models with heterogeneous reservation utilities, that is, $\partial\left[F(\theta)/f(\theta)\right]/\partial\theta \geq 0$. It guarantees that the seller has an incentive to discriminate across consumers.[18] Consider next the condition on $\overline{u}''(\theta)$. This condition ensures that consumers whose IR constraints bind are offered quantity discounts. The convexity of $\overline{u}(\cdot)$ implies that outside consumption opportunities are increasingly more valuable for consumers of higher types. Since $\overline{u}'(\theta) = \nu(\overline{q}(\theta))$ by assumption (H), by offering larger quantities at lower marginal prices, the seller can satisfy higher types' IR constraints and induce them to buy more than lower types, thereby separating them from lower types. Quantity discounts are then optimal for the seller.

By comparing the first-order condition in (9) with that for the first-best allocation, $T'(q(\theta)) = c$, it is immediate that the quantity provided to a consumer of type θ is below the first best when $\gamma(\theta) > F(\theta)$ but above the first best when $\gamma(\theta) < F(\theta)$. Underprovision arises when the reservation utility for higher consumer types is close enough to that for lower types that participation constraints tend to bind for lower rather than for higher types. In this case, as in the standard model, higher types have an incentive to imitate the behavior of lower types. But since higher types enjoy a higher marginal benefit from consuming the good, a seller can separate higher types from lower ones by decreasing the offered quantities meant for lower types below lower types' first-best level of consumption. This way, a seller makes the purchase of small quantities unattractive to higher types.

Overprovision, instead, arises when the reservation utility for higher consumer types is larger enough than that for lower types that participation constraints tend to bind for higher types. In this case, a seller needs to induce higher types to buy in the first place and can do so while separating them from lower types by offering quantities meant for higher types that are larger than the first best at marginal prices below marginal cost. By doing so, a seller not only can induce higher types to participate by purchasing these large quantities but also can distinguish them from lower types, who naturally prefer smaller quantities. The seller can differentiate consumers because lower types would need to purchase much larger quantities than desirable to them to imitate the behavior of higher types, that is, quantities above the first-best level of consumption of higher types.[19]

Nonlinear versus Linear Pricing. A natural question is whether consumers are better off under nonlinear or linear pricing. Under linear pricing, a seller charges the unit price p_m for any quantity provided. A consumer of type θ chooses the quantity $q_m(\theta)$ and obtains utility $u_m(\theta) = \theta\nu(q_m(\theta)) - p_m q_m(\theta)$ from purchasing the good. It turns out that when all consumers participate under both pricing schemes and nonlinear

[18] All these conditions on the distribution of types are satisfied, for instance, by a uniform distribution with $\underline{\theta} \geq 0$ and a four-parameter beta distribution with $a \geq 1$ and $b = 1$.

[19] Since $\gamma(\theta) \leq 1$, the augmented model gives rise to higher levels of consumption and, correspondingly, lower marginal prices relative to the standard model. Given that higher quantities may be offered at a higher price $T(q)$, the overall effect on consumers' utility is ambiguous. When $\overline{u}(\underline{\theta}) \geq \overline{u}$ so that the reservation utility in the augmented model, $\overline{u}(\theta)$, is higher than in the standard model, $\overline{u}$, for each type, consumer surplus is clearly higher in the augmented model.

pricing entails quantity discounts, consumers tend to be better off under linear pricing. Intuitively, linear pricing is preferred when the quantity provided under linear pricing is larger: nonlinear pricing just allows a seller to better extract consumer surplus. Perhaps surprisingly, consumers prefer linear pricing even when the quantity provided under linear pricing is smaller. In this case, a seller who can price discriminate tends to charge high prices for the greater quantity provided. Critically, however, a consumer who is excluded from the market under linear pricing but included under nonlinear pricing prefers nonlinear pricing. For instance, consumers who have access to generous enough outside consumption possibilities in that $\overline{q}(\theta)>q^{FB}(\theta)$ can be excluded under linear pricing and so are better off under nonlinear pricing.

Proposition 3 (*Nonlinear versus Linear Pricing*)**.** *The following results hold:*
1) Suppose that (FP) holds 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 $q=q(\theta)$, $q(\theta)>q_m(\theta)$, and $\gamma(\theta)<1$, then a consumer of type θ is better off under linear pricing.
2) Suppose that $\nu''(\cdot)<0$, $s(\theta,\overline{q}(\theta))\geq\overline{u}(\theta)$, and $\overline{q}(\theta)>q^{FB}(\theta)$ for consumer types in the interval $[\theta',\theta'']$. If there exists a type $\hat{\theta}$ in $[\theta',\theta'']$ with $u_m(\hat{\theta})=\overline{u}(\hat{\theta})$, then an interval of consumer types in $[\hat{\theta},\theta'']$ is excluded under linear pricing, but these consumer types participate under nonlinear pricing and so are better off under nonlinear pricing.

To understand the condition $\overline{q}(\theta)>q^{FB}(\theta)$ in part 2) of Proposition 3 , note that since $\overline{u}'(\theta)=\nu(\overline{q}(\theta))$ by assumption (H) or (BCH), large values of $\overline{q}(\theta)$ are associated with a rapidly increasing reservation utility profile.[20] To induce consumers with 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 price would lower profits from all existing consumers for the benefit of including only a few more. Hence, it would not be profitable to include such consumers under linear pricing. Proposition 3 then highlights a dimension along which nonlinear pricing may be more efficient than linear pricing. Whenever different consumers can be charged different prices, a seller may have an incentive to serve those consumers who would be unprofitable under linear pricing; see the Supplementary Appendix for an example. We examine the extent to which this implication of our model is borne out in the data in Section 5.3.

Income Transfers. Here we show that when consumers face a budget constraint for the seller's good, income transfers increase consumption but also typically lead to an increase in prices, as the seller adjusts offered quantities and prices in response to consumers' greater ability to pay. Intuitively, when consumers are constrained by a budget for the seller's good, an increase in their income affects prices by creating an incentive for the seller to extract more surplus. For instance, suppose 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: as 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 con-

[20]Note that Proposition 3 requires the existence of a type at risk of exclusion, namely, a type whose utility equals $\overline{u}(\theta)$ under linear pricing. Also, $\overline{q}(\theta)>q^{FB}(\theta)$ cannot arise in general when $\gamma(\theta)=1$ for all types, as in the standard model, since $q(\theta)\leq q^{FB}(\theta)$ for all types in this case. See Corollary 1 in Jullien (2000) for a proof that if $\overline{q}(\theta)\geq q^{FB}(\theta)$ for all types, then $q(\theta)\geq q^{FB}(\theta)$.

sumers' behavior. Indeed, the seller maximizes profits by increasing the price $T(q)$ of each quantity by the amount of the transfer.

Now consider the more interesting case in which the transfer depends on consumers' characteristics. In the villages we study, the Progresa transfer depends on a household's income and number of children. Given that poorer households tend to have more children, transfers are larger for poorer households and thus are effectively progressive in income; see Attanasio et al. (2013). For the normal goods we consider, since poorer households consume less than richer households and our model implies a monotone relationship between types and quantities, assuming that the transfer satisfies $\tau'(\theta) \leq 0$ seems then consistent with the data.

To understand the impact of such a transfer, recall that a consumer of type θ pays $t(\theta)$ to purchase $q(\theta)$ from the seller and the rest of her income to purchase z subject to the subsistence constraint $z = Y - t(\theta) \geq \underline{z}(\theta, q(\theta))$ before the transfer is introduced. Thus, the consumer's budget constraint for the seller's good is $t(\theta) \leq Y - \underline{z}(\theta, q(\theta))$ for the menu pair $(t(\theta), q(\theta))$. Once the consumer receives the transfer $\tau(\theta)$, her ability to pay correspondingly increases and her budget constraint for the seller's good becomes $t_\tau(\theta) \leq Y + \tau(\theta) - \underline{z}(\theta, q_\tau(\theta))$ for the menu pair $(t_\tau(\theta), q_\tau(\theta))$. As in the case of a uniform transfer, the seller can then ask for a higher price without excluding any consumer. But since consumers' ability to pay increases differentially with the transfer, the seller will typically charge different consumer types different prices. In particular, it turns out that since $\tau'(\theta) \leq 0$, the marginal price for a consumer who spends her entire budget on the seller's good before and after the transfer must decrease. Thus, any such consumer demands a higher quantity. To preserve incentive compatibility, a seller must then offer higher quantities, and thus lower marginal prices, to other consumers as well. As a result, consumption increases and the marginal price paid decreases for at least some consumers. Although the marginal price $T'(q)$ decreases, the price $T(q)$ paid by any such consumer increases by an argument analogous to that in the case of a uniform transfer.

Proposition 4 (*Income Transfers*)**.** *Let $\nu''(\cdot) < 0$. Consider a transfer such that $\tau'(\theta) \leq 0$ with strict inequality for at least an interval of consumer types. Suppose that the budget constraint binds before and after the transfer for a consumer of type θ' in such an interval. Then, under the conditions of Proposition 1 with $I_{\theta q}(\theta, q) \geq 0$, the transfer leads to greater consumption and a higher price schedule $T(q)$ with lower marginal prices $T'(q)$ for all types in some interval $[\theta_{\min}, \theta_{\max}]$ including θ'.*

As both $T(q)$ and q increase for some consumer types, the effect of the transfer on the unit price of the good, $T(q)/q$, and so on the intensity of price discrimination is ambiguous. We now show, however, that the intensity of price discrimination, as measured by the size of quantity discounts (the absolute value of $T''(q)$), increases with the transfer for at least some consumers when the distributions of quantities purchased before and after the transfer can be ranked.

To examine the impact of the transfer on the nonlinearity of prices, we compare the curvature of the price schedule before and after the transfer type by type or, equivalently, at the same percentiles in the distributions of quantities before and after the transfer is introduced. Since optimal quantity profiles are monotone, a given percentile in the two quantity distributions corresponds to the same type. Formally, denote by $\{q_\tau(\theta)\}$

the quantity profile after the transfer is introduced and by $G(q)$ and $G_\tau(q)$, respectively, the cumulative distribution functions of quantities before and after the transfer with associated probability density functions $g(q)$ and $g_\tau(q)$. For any quantity $q = q(\theta)$ purchased before the transfer by a consumer of type θ, the quantity $q_\tau = q_\tau(\theta)$ purchased by the consumer after the transfer satisfies

$$G_\tau(q_\tau) = G_\tau(q_\tau(\theta)) = \Pr(\tilde{\theta} \leq q_\tau^{-1}(q_\tau) = \theta) = F(\theta) = \Pr(\tilde{\theta} \leq q^{-1}(q) = \theta) = G(q(\theta)) = G(q), \quad (10)$$

where $\tilde{q}_\tau = q_\tau(\tilde{\theta})$ and $\tilde{q} = q(\tilde{\theta})$. Here we consider a transfer that increases consumption such that

$$g_\tau[G_\tau^{-1}(t)] \leq g[G^{-1}(t)] \text{ up to some } t \in (G_\tau(0), 1). \quad (11)$$

Condition (11) states that up to a certain percentile $q_{\max}$ in the distribution of quantities before and after the transfer, the probability density function of quantities purchased *after* the transfer is smaller than the probability density function of quantities purchased *before* the transfer at each percentile. Hence, the distribution $G_\tau(\cdot)$ assigns more mass to larger quantities than the distribution $G(\cdot)$ up to t. Indeed, if (11) applies to all $t \in (G_\tau(0), 1)$, then $G_\tau(\cdot)$ first-order stochastically dominates $G(\cdot)$; see Dharmadhikari and Joag-Dev (1983). Note that if the transfer leads to greater consumption by all types, then $G_\tau(\cdot)$ first-order stochastically dominates $G(\cdot)$. Intuitively, when $q_\tau(\theta) \geq q(\theta)$, a given percentile in the distribution of purchased quantities after the transfer corresponds to a larger quantity than before the transfer. Therefore, condition (11) simply amounts to a strengthening of the dominance ordering between $G_\tau(\cdot)$ and $G(\cdot)$ implied by the transfer up to some probability t.

Corollary 1 (*Income Transfers and Price Discrimination*)**.** *If $\nu'''(\cdot) \leq 0$ and the transfer $\tau(\theta)$ leads to greater consumption for all consumers so that (11) holds, then there exists a percentile $q_{\max}$ in the distribution of quantities before and after the transfer such that $T''_\tau(q_\tau) \leq T''(q)$ at all percentiles up to $q_{\max}$.*

When $T''(q) \leq 0$, the transfer then leads to greater price discrimination in the form of larger discounts. In Section 5.4, we show that these implications of our model are supported by the data.[21]

4 Identification and Estimation

In this section, we discuss the identification and estimation of the model's primitives, which builds on intuitions from Perrigne and Vuong (2010). Intuitively, the pricing behavior of a seller depends on the distribution of consumer types in a village. Since this distribution can be mapped into that of purchased quantities, the distribution of consumer types can be recovered from the joint distribution of observed prices and quantities. Although the model's primitives can be identified and estimated semiparametrically based on this intuition, here we derive estimators that use flexible parametric functions, partially to accommodate the sparsity of the data in some villages. We estimate the model using data for three commodities—rice, kidney beans, and

[21] See the Supplementary Appendix for an example in which $\nu(q)$ is a HARA (hyperbolic absolute risk aversion) function and the intensity of price discrimination increases for some consumers in response to a transfer.

sugar—that we chose for three reasons. First, as discussed in Section 2, they are commonly consumed, so the full participation assumption is likely to be valid and we observe a large number of transactions for them. Second, they are goods of homogeneous quality. Hence, the variation in prices across quantities we document is likely to reflect just quantity discounts. Third, they are normal goods whose consumption increases with income. Thus, assuming that households' marginal willingness to pay and absolute ability to pay are related, as we do, is plausible. See Appendix B for omitted details and proofs.

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 $c'(Q)$ at the total quantity provided $Q = \int_{\underline{\theta}}^{\overline{\theta}} q(\theta) f(\theta) d\theta$, and the determinants of participation in the market, namely, 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.[22] Under standard assumptions, we show that these primitives are identified in each village from data on consumers' expenditures and purchases, which provide information, respectively, about $T(q)$ and q. Note that $\overline{u}(\theta)$ and $\Upsilon(\theta)$ are identified only for households with binding IR or BC constraints.[23] We refer to the cumulative multiplier $\gamma(\theta)$ associated with the IR or BC constraints 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, where $s(\theta, \overline{q}(\theta))$ is the social surplus at the reservation quantity $\overline{q}(\theta)$: it states that a seller obtains nonnegative profits from each consumer's type at the quantity $\overline{q}(\theta)$. This approach is justified by the fact that the overwhelming majority of households in each village purchase the three goods we focus on, namely, rice, kidney beans, and sugar, as discussed. We also adopt the normalization $\underline{\theta} = 1$, since a scaling assumption is required for identification. We denote by $G(q)$ the cumulative distribution function of quantities purchased in a village and by $g(q)$ the associated probability density function. Since $G(q)$, $g(q)$, the price schedule $T(q)$, and its derivatives are identifiable from information on prices and quantities in our data, 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 optimal choice of quantity. We use this latter condition to identify a seller's cost structure. By relying exclusively on information on prices and quantities, though, we can identify a seller's marginal cost only at the total quantity of a good provided in a village, Q. Nonetheless, based on this information alone, we can identify all primitives up to consumers' coefficient of absolute risk

[22] Consumers' marginal willingness and absolute ability to pay, here θ and w, are highly correlated in the data, as implied by the strong relationship between consumption and income and the fact that the commodities we consider are normal goods.

[23] 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 $\overline{u}(\theta)$ or $\Upsilon(\theta)$ on Θ' and are appropriately adjusted for the remaining types.

aversion under the assumption that $v(\theta, q) = \theta\nu(q)$, which we maintain from now on.[24] This specification of utility is ubiquitous in the literature on auctions and nonlinear pricing for its tractability (see Guerre et al. (2000) and Perrigne and Vuong (2010)), so 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, denote by $\underline{q} \equiv q(\underline{\theta})$ and $\overline{q} \equiv q(\overline{\theta})$ the smallest and largest observed quantities of a good purchased in a village. Recall that since $q(\theta)$ is an increasing function, it follows that $G(q) = \Pr(\tilde{q} \leq q) = \Pr(\tilde{\theta} \leq q^{-1}(q) = \theta) = F(\theta)$; that is, the cumulative distribution function of types is identified by that of quantities, and so $f(\theta) = g(q)q'(\theta)$. 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 thus the set of consumers whose participation (or budget) constraints bind. Formally, rewrite (9) as

$$\frac{g(q)}{\varphi(q)} \left[\frac{c'(Q)}{T'(q)} - 1 \right] = G(q) - \gamma(\theta(q)), \tag{12}$$

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, $A(q) = -\nu''(q)/\nu'(q)$. As a preliminary step, we argue that $c'(Q)$ is identified up to the ratio $\varphi(\overline{q})/\varphi(\underline{q})$. To this purpose, it is easy to show that taking derivatives of both sides of (12) and integrating the resulting expressions from $\underline{q}$ and $\overline{q}$ yields that

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

see the proof of Proposition 5 in Appendix A. Since $g(q)$ and $T'(q)$ are identified, $c'(Q)$ is then identified up to $\varphi(\overline{q})/\varphi(\underline{q})$. Now, 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 (12) also implies that

$$\gamma(\theta(q)) = G(q) + g(q) \left[1 - \frac{c'(Q)}{T'(q)} \right] \left[\frac{T''(q)}{T'(q)} + A(q) \right]^{-1}. \tag{13}$$

Thus, 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. But since $\varphi(q)/\varphi(q)$ only depends on $\theta'(q)/\theta(q)$, which is known up to $A(q)$, it follows that $c'(Q)$ is also identified if $A(q)$ is known. Hence, $\gamma(\theta(q))$ is identified just up to $A(q)$.

Equation (13) clarifies that the identification of $c'(Q)$ and the multiplier $\gamma(\theta(q))$ requires some knowledge of the shape of the utility function. In estimation, we circumvent this issue by specifying $\gamma(\theta(q))$ as a flexible parametric function of q so that we only need to estimate $c'(Q)$ and the parameters of this function.

[24] 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 may not be 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 and q, $\gamma(\theta)$ 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.

Proposition 5. *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 $T''(q)$.*

Note that when the multiplier is constant on $[\underline{\theta},\overline{\theta})$, $\gamma(\theta(q))=\gamma$ equals $G(q)$ only at one quantity in $[\underline{q},\overline{q})$. In this case, the constant γ is identified by the value of $G(q)$ at the quantity at which $T'(q)$ equals $c'(Q)$ by (12), where $c'(Q)$ is identified from $A(q)$ as discussed.[25]

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 (12) 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

$$\log(\theta(q)) = \log(\theta(\underline{q})) + \int_{\underline{q}}^{q} \frac{\partial \log(\theta(x))}{\partial x} dx = \log(\theta(\underline{q})) + \int_{\underline{q}}^{q} \frac{g(x)[T'(x) - c'(Q)]}{T'(x)[\gamma(\theta(x)) - G(x)]} dx, \qquad (14)$$

where the first equality in (14) follows by integrating $\partial\log(\theta(q))/\partial q$ with respect to quantity.[26] Once $c'(Q)$ and $\gamma(\theta(q))$ are identified, expression (14) implies that $\theta(q)$ is identified as well 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 $F(\theta)=G(q)$, as argued.

Proposition 6. *In a village, the support of consumers' marginal willingness to pay $\theta(q)$ is identified from the cumulative distribution function $G(q)$ and the probability density function of quantities $g(q)$, 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 and the first derivative of $\theta(q)$.*

Utility Function and Schedule of Reservation Utility. Note that knowledge of the coefficient of absolute risk aversion alone implies that the base marginal utility function $\nu'(q)$ is identified up to a level normalization. Once the marginal price schedule and the type support are identified, though, $\nu'(q)$ is identified even without such a normalization from $T'(q)$ and $\theta(q)$ by the incentive compatibility condition $\theta(q)\nu'(q)=T'(q)$. Then, we can recover $\nu(q)$ from $\nu'(q)$ up to one point, say, its value at $q'=q(\theta')$, as $\nu(q)=\nu(q')-\int_q^{q'}\nu'(x)dx$ 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 $\overline{u}(\theta)=\theta\nu(q(\theta))-T(q(\theta))$.

Proposition 7. *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 up*

[25] When the standard model is *known* to apply, knowledge of $A(q)$ is unnecessary since $\gamma(\theta(q))$ equals one at all quantities, so $c'(Q)$ is identified by $T'(q)$ at the largest quantity.

[26] In (14), $\gamma(\theta(q))$ is interpreted as a function of q. The integrand 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 (12). It is well defined when $\gamma(\theta(q))=G(q)$ and $T'(q)=c'(Q)$ as long as the slope of $\gamma(\theta(q))$ differs from $g(q)$ at such quantities. Specifically, denoting by q^s any quantity such that $\gamma(\theta(q))=G(q)$ and $T'(q)=c'(Q)$, note that the limit of the integrand as q converges to q^s is $g(q^s)T''(q^s)/\{T'(q^s)[\gamma'(\theta(q^s))\theta'(q^s) - g(q^s)]\}$. That $\gamma'(\theta(q^s))\theta'(q^s)$ in general differs from $g(q^s)$ is 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 a level normalization. The reservation utility (or budget) function is identified for all consumers whose participation (or budget) constraints bind, that is, all consumers with $\partial\gamma(\theta(q))/\partial q > 0$.

4.2 Estimation

We estimate the model separately in each village for each of the three goods considered in two steps. In the first step, we parameterize the functions $T(q)$, $G(q)$, and $\gamma(\theta(q))$ and estimate their parameters by maximum likelihood together with the model's primitives $c'(Q)$, $\theta(q)$, and $\nu'(q)$. The assumed expressions for $T(q)$ and $G(q)$ and a seller's first-order condition provide the three estimating equations for the parameters of $T(q)$, $G(q)$, $\gamma(\theta(q))$, and for $c'(Q)$. We estimate $\theta(q)$ and $\nu'(q)$ in each village as known transformations of $T'(q)$, $G(q)$, $\gamma(\theta(q))$, and $c'(Q)$ based, respectively, on equation (14) and the local incentive compatibility condition $\nu'(q) = T'(q)/\theta(q)$. In the second step, we estimate $f(\theta)$ in each village from the estimated $\theta(q)$ via a kernel density estimator.[27]

Price Schedule and Distribution of Quantities. Our data contain information on the quantities purchased and the prices paid in each village for each good we study, from which unit prices can easily be computed. Denote by N_{vj} the number of households purchasing good j in village v and by q_{vji} the quantity of the good purchased by household i. We estimate the price schedule of good j in village v as

$$\log[T_{vj}(q_{vji})] = t_{vj0} + t_{vj1}\log(q_{vji}) + \varepsilon^{p}_{vji}, \tag{15}$$

where $T_{vj}(q_{vji}) \equiv E[p(q_{vji})|q_{vji}]q_{vji}$, $p(q_{vji})$ is the unit price of quantity q_{vji}, and ε^{p}_{vji} is measurement error. The assumption implicit in (15) is that expenditure, and so unit values, rather than quantities are contaminated by error. We use the mean unit value $E[p(q_{vji})|q_{vji}]$ of quantity q_{vji} to construct $T_{vj}(q_{vji})$ to minimize the impact of measurement error in unit values due, for instance, to recall or recording error as well as for consistency with our model. Namely, although multiple unit values may be associated with a given quantity in a village, our model implies that the price schedule is a function of quantity rather than a correspondence. We treat quantity as exogenous since the quantities purchased and prices paid provide direct information on the price schedule in a village and the price schedule is a deterministic function of quantity in our model.

We parameterize the cumulative distribution function of the quantities of good j purchased in village v as a logistic function with index $\Phi_{vj}(\cdot)$,

$$G_{vj}(q_{vji}) = \frac{\exp\{\Phi_{vj}(q_{vji}) + \varepsilon^{g}_{vji}\}}{1 + \exp\{\Phi_{vj}(q_{vji}) + \varepsilon^{g}_{vji}\}}, \tag{16}$$

where $G_{vj}(q_{vji})$ is the empirical cumulative distribution function of purchased quantities and $\Phi_{vj}(\cdot)$ is a flexible polynomial (up to the third degree, including a fractional polynomial). In each village, we select the specification of $\Phi_{vj}(\cdot)$ with the lowest value of the Akaike information criterion (AIC). Note that ε^{g}_{vji} in (16)

[27] Since the convergence of the parametric estimator of $\theta(q)$ is faster than that of the nonparametric estimator of $f(\theta)$, the estimation of $\theta(q)$ does not affect this second step. Note that if $f(\theta)$ is interpreted as the probability mass function associated with the empirical $G(q)$, this second step is unnecessary.

captures not only recall or recording error but also error in observed purchase frequencies resulting from the timing of the Progresa interview. For instance, in the week preceding the interview, a household may not have purchased a good it commonly buys and so may be assigned no recorded purchase. In general, such an error may lead to understating or overstating the fraction of households purchasing a particular quantity.

Marginal Cost and Multiplier. By (12), we can relate the cumulative distribution function of quantities purchased $G(q)$ to the marginal cost $c'(Q)$, the multiplier $\gamma(\theta(q))$, and the unit price $p(q)$ as

$$G(q) = \left[\frac{1}{T'(q)} - \frac{1}{c'(Q)}\right] x(q) + \gamma(\theta(q)) = \left[\frac{1}{t_1 p(q)} - \frac{1}{c'(Q)}\right] x(q) + \gamma(\theta(q)), \tag{17}$$

where $x(q) \equiv c'(Q)g(q)\theta(q)/\theta'(q) > 0$ and the second equality in (17) follows by (15), which implies that the unit price $p(q) = T(q)/q$ can be expressed as $p(q) = T'(q)/t_1$. Denote the marginal cost of the total quantity of good j purchased in village v by $c'_{vj}(Q_{vj})$. We specify the auxiliary function $x(\cdot)$ as a positive function with up to two parameters, $x_{vj}(q_{vji}) = \chi_{vj0} + \chi_{vj1} q_{vji}$; given the limited granularity of our data, estimating $x(q)$ more flexibly would be infeasible. Since the multiplier has the properties of a cumulative distribution function, we estimate it as $\gamma_{vj}(q_{vji}) = \exp\{\Gamma_{vj}(q_{vji})\}/(1 + \exp\{\Gamma_{vj}(q_{vji})\})$ for good j in village v, where the index $\Gamma_{vj}(\cdot)$ is a polynomial (up to the second degree). In each village, we select the specification of $x_{vj}(\cdot)$ and $\Gamma_{vj}(\cdot)$ with the lowest AIC value. Then, expression (17) becomes

$$\begin{aligned} G_{vj}(q_{vji}) &= \left[\frac{1}{p_{vj}(q_{vji})} - \frac{1}{\underline{c}'_{vj}(Q_{vj})}\right] (\underline{\chi}_{vj0} + \underline{\chi}_{vj1} q_{vji}) + \gamma_{vj}(q_{vji}) + \varepsilon^s_{vji} \\ &= -\frac{\underline{\chi}_{vj0}}{\underline{c}'_{vj}(Q_{vj})} + \underline{\chi}_{vj0}\frac{1}{p_{vj}(q_{vji})} - \frac{\underline{\chi}_{vj1}}{\underline{c}'_{vj}(Q_{vj})} q_{vji} + \underline{\chi}_{vj1}\frac{q_{vji}}{p_{vj}(q_{vji})} + \frac{\exp\{\Gamma_{vj}(q_{vji})\}}{1 + \exp\{\Gamma_{vj}(q_{vji})\}} + \varepsilon^s_{vji}, \end{aligned} \tag{18}$$

which is the sum of a linear-in-parameters function, given by the first four terms, and a nonlinear one, where $\underline{c}'_{vj}(Q_{vj}) \equiv c'_{vj}(Q_{vj})/t_{vj1}$, $\underline{\chi}_{vj0} \equiv \chi_{vj0}/t_{vj1}$, and $\underline{\chi}_{vj1} \equiv \chi_{vj1}/t_{vj1}$, and ε^s_{vji} is measurement error.[28]

Support of Consumer Types and Utility Function. We normalize $\underline{\theta}$ to 1 and specify households' (log) marginal willingness to pay in village v for good j by (14) 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 households purchasing the good, Q_{vj} is the total quantity purchased in village v, $T'_{vj}(q_{vji})$ is computed from (15), and $\gamma_{vj}(q_{vji})$ and $G_{vj}(q_{vji})$ are specified as discussed.[29] Using the local

[28]To see how the parameters of (18) are identified, note that there exists at least one quantity q^{FB}_{vj} such that $\gamma(\theta(q^{FB}_{vj})) = G_{vj}(q^{FB}_{vj})$; the mean unit price of this quantity identifies $\underline{c}'_{vj}(Q_{vj})$. With the multiplier at the largest quantity equal to one, as long as the first (or any higher) derivative of the inverse mean price schedule with respect to quantity evaluated at the largest quantity is zero and the first (or any higher) derivative of the multiplier function with respect to quantity evaluated at the smallest quantity is zero, then $\underline{\chi}_{vj0}$ and $\underline{\chi}_{vj1}$ are identified. Once $\underline{c}'_{vj}(Q_{vj})$, $\underline{\chi}_{vj0}$, and $\underline{\chi}_{vj1}$ are identified, the parameters of $\gamma_{vj}(\cdot)$ are identified by (18) evaluated at up to two more quantities in addition to q^{FB}_{vj}.

[29]In a slight abuse of notation, we denote the derivative of the exponential of the predicted log tariff by $T'_{vj}(\cdot)$.

incentive compatibility condition $\nu'(q) = T'(q)/\theta(q)$ and the form of $\theta(q)$, we estimate marginal utility as

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

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 first discuss our sample selection criteria, present the estimates of the model's primitives, and show the fit of the model to the data. Since we consider many villages, we graphically represent the point estimates of the objects of interest and report their associated t-statistics in Appendix B. We then use the model to analyze the distortions implied by the price discrimination we observe and evaluate the impact 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 purchased in each village. We use this reduced form to estimate the effect of the Progresa transfers on the prices of each good by exploiting the experimental variation in the data induced by the introduction of the transfers in a randomly selected subset of villages. Based on this reduced form, we then evaluate the ability of our model to account for the impact of Progresa on prices. See Appendix B and the Supplementary Appendix for omitted results and details.

5.1 Estimation Sample

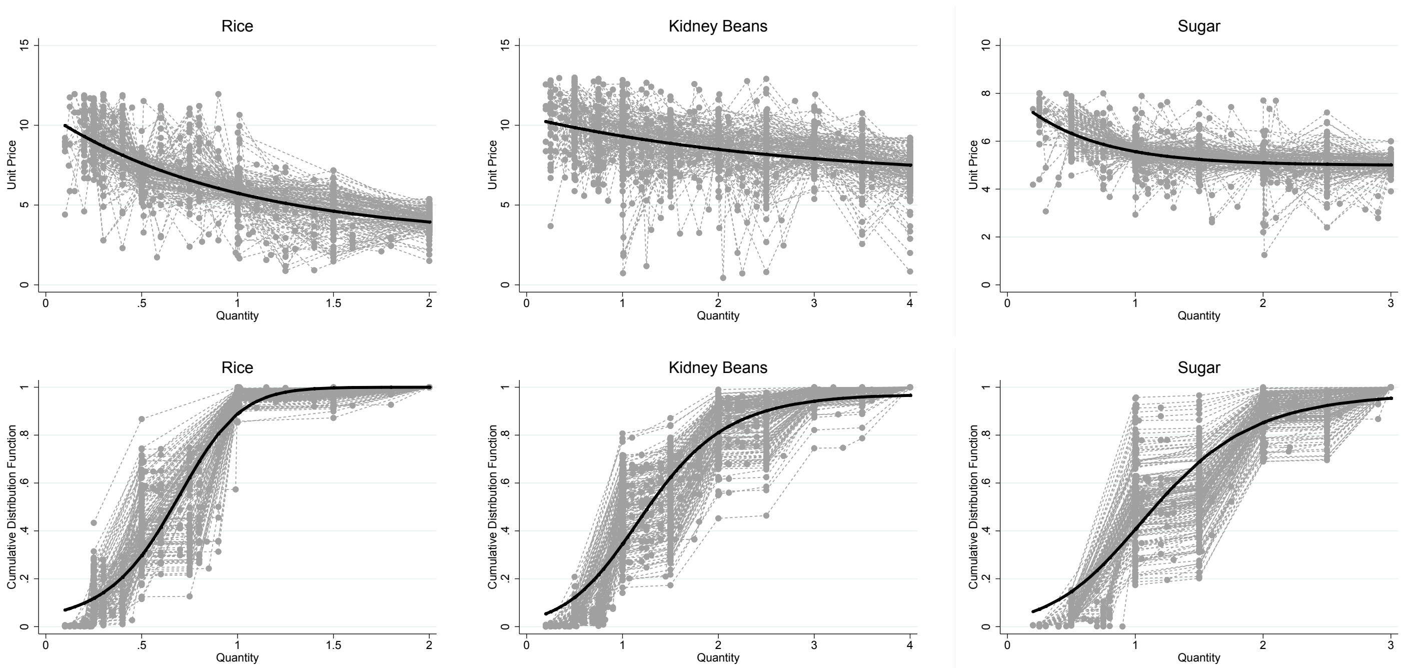
Here we describe our sample selection criteria and present key statistics from the resulting estimation sample.

Sample Selection. We use five waves of the Progresa evaluation surveys, namely, October 1998, May and November 1999, November 2000, and 2003, for rice, kidney beans, and sugar. Although it might be natural to define a village and so the relevant market for a good at the level of a Mexican locality, here we define a village as a Mexican municipality, as discussed. However, estimates of the model based on villages defined as localities, which we report in the Supplementary Appendix, are very similar to those based on villages defined as municipalities. To minimize the impact of measurement error, we ignore purchases reported in units different from kilos and exclude extreme observations—we drop the top 5% of quantities and expenditures, the latter expressed relative to their level in October 1998, and trim the top 1% of the resulting mean unit prices in each village. We focus on villages where at least 50% of the unit prices decline with quantity and with at least 75 observations on each good of interest. These restrictions imply the loss of only a few villages: the original sample of 191 municipalities is reduced to 174 for rice, 183 for kidney beans, and 185 for sugar.[30]

[30]Focusing on villages with at least 50% of the unit prices declining with quantity accounts for a small loss of villages, which is primarily due to the irregularity of price schedules in those villages.

Prices and Quantities. In the top panels of Figure 1, we report the schedule of mean unit prices per quantity in each village for each good computed as explained above (together with an interpolating solid line). Similarly, in the bottom panels, we report the cumulative distribution function of quantities purchased in each village. In most villages, the unit price of each good declines with quantity, which implies that 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 by comparing the top and bottom panels of the figure. Thus, most households are affected by the nonlinearity of prices and face significant quantity discounts. For instance, the mean unit price of the smallest quantity of rice, 0.1 kilos, is on average more than 8 pesos, whereas the unit price of the largest quantity, 2 kilos, can be as low as 1.5 pesos.

Figure 1: Unit Prices and Cumulative Distribution Function of Quantities



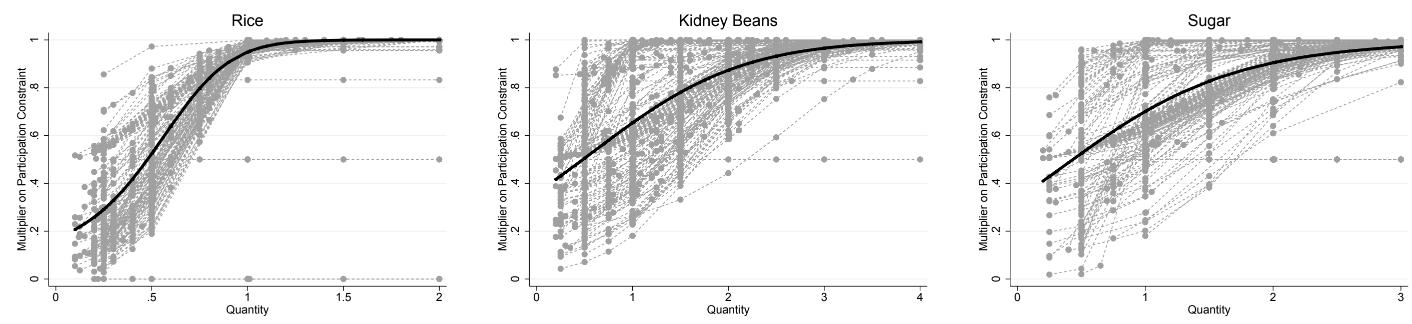
5.2 Estimation Results

The core elements of our model are the multiplier on participation (or budget) constraints, the distribution of consumer types, and consumers' utility function. In this subsection, we present their estimates based on the sample of municipalities and illustrate the fit of the model to the data; see Appendix B for the estimates of marginal cost and the omitted t-statistics of all estimates. We successfully estimate the model for 173, 183, and 185 of the 174, 183, and 185 municipalities in the estimation samples for rice, kidney beans, and sugar, respectively.[31] The analogous figures for localities are 363, 408, and 451 of the 368, 411, and 453 localities in the estimation samples for the three goods that satisfy the same selection rules applied to municipalities.

[31] In some villages, although $\log(\theta(q))$ is estimated, $\theta(q)$ is recorded as missing when its value is exceedingly large, so $f(\theta)$ is not estimated at any such support point.

Estimates of Multipliers. Figure 2 reports the estimated multipliers on participation (or budget) constraints, $\gamma(\theta(q))$, for each quantity purchased in each village by good. Recall that, by construction, the multiplier ranges between 0 and 1. We estimate that its mean across quantities and villages is 0.707 for rice with a standard deviation of 0.323; 0.789 for kidney beans with a standard deviation of 0.237; and 0.798 for sugar with a standard deviation of 0.218. For each good, the multiplier varies substantially across quantities and is smaller than one for most of them: the 25th, 50th, and 75th percentiles in the distribution of the estimated $\gamma(\theta(q))$ across quantities and villages are, respectively, 0.435, 0.868, and 0.998 for rice; 0.603, 0.897, and 0.992 for kidney beans; and 0.641, 0.892, and 0.982 for sugar.

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



As discussed, the shape of the multiplier function distinguishes different instances of our model. Note that the multiplier is estimated to be constant only in a handful of villages. Based on tests of the individual and joint significance of the parameters of $\gamma(\theta(q))$, we reject the hypothesis that the standard model applies, that is, $\gamma(\theta(q))=1$ at all q, except for one village for kidney beans.[32] For an intuition about why villages do not conform to the standard model, recall that the seller's first-order condition can be expressed as in (17). For the multiplier to be constant, the term in brackets should replicate the variability of $G(q)$, since $x(q)$ is positive and estimated to be roughly constant over the quantities that most households buy. Thus, $p(q)$ and $G(q)$ should be approximately inversely related. Indeed, as Figure 1 shows, the unit price schedule $p(q)$ of each good starts on average at a high value and is approximately decreasing, whereas $G(q)$ starts at a low value and is increasing. But an important departure from this inverse relationship between $p(q)$ and $G(q)$ is that whereas the curvature of $p(q)$ tends to be most pronounced at *small* quantities, that of $G(q)$ is most pronounced at *intermediate* quantities. This difference in the shapes of $p(q)$ and $G(q)$ is accommodated by $\gamma(\theta(q))$ varying across quantities.

Estimates of Type Distribution and Marginal Utility. In the top panels of Figure 3, we report the estimates of base marginal utility, $\nu'(q)$, and in the bottom panels of the figure, we report the estimates of marginal utility, $\theta(q)\nu'(q)$, for each quantity purchased in each village by good. Note that $\nu'(q)$ decreases with quantity in all villages, as consistent with the model, although no such monotonicity restriction has been

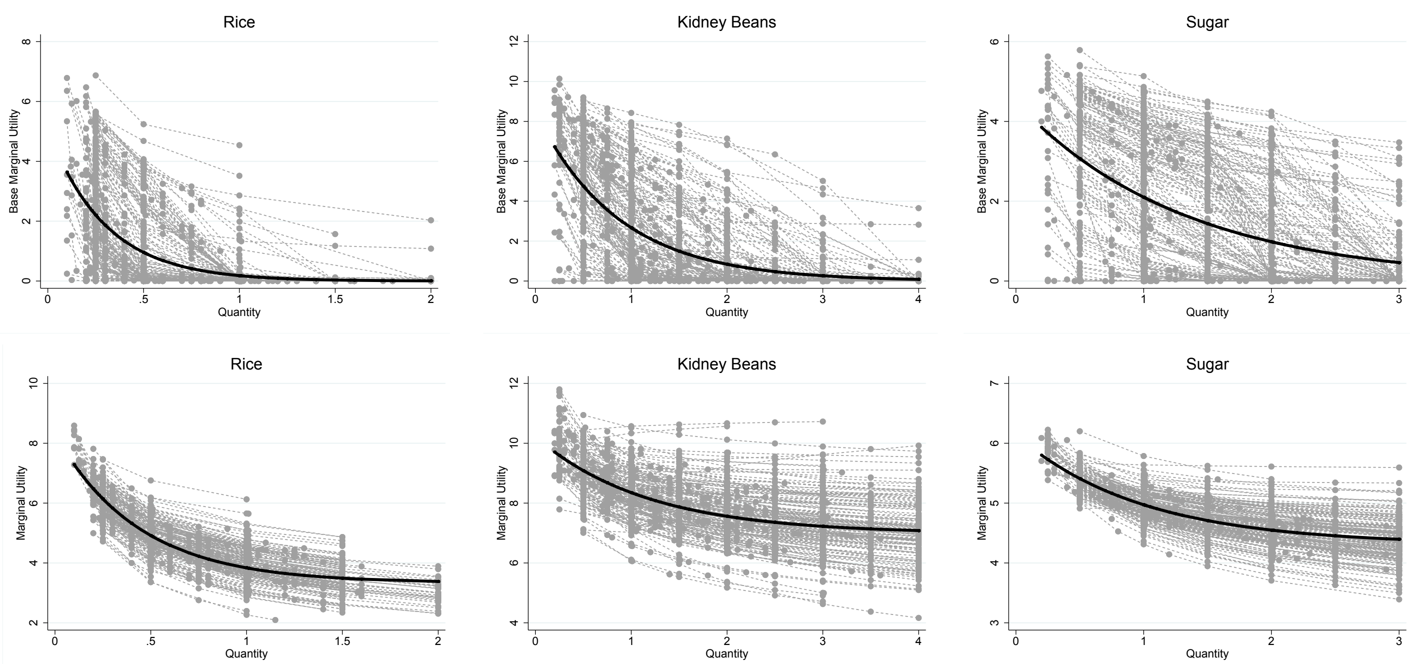
[32]We perform significance tests on the estimated value of the multiplier function $\gamma(\theta(q))$ in each village for each good so as to determine whether the multiplier is ever constant across quantities and, if it is constant across all quantities, whether it is significantly different from zero and one. We then construct the predicted multiplier for each quantity accordingly. Note that if no parameter of $\gamma_{vj}(\cdot)$ is significant, then $\gamma(\theta(q))$ equals 0.5 at all quantities.

Table 2: Distribution of Log Consumer Types

Good	Percentiles of Log Consumer Types										
	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%
Rice	0.4	0.6	2.0	3.0	4.9	7.2	11.3	23.1	42.3	104.5	784.7
Kidney Beans	0.2	0.4	0.6	1.0	1.7	2.7	4.4	7.5	15.3	52.4	93.8
Sugar	0.2	0.2	0.3	0.5	0.8	1.2	2.1	4.3	11.6	24.4	57.8

imposed in estimation. In nearly all villages, instead, consumers' estimated marginal willingness to pay $\theta(q)$ increases with quantity, as consistent with the incentive compatibility condition of our model, which we have not imposed, and rapidly so at large quantities. Indeed, the estimated support of consumer types for each good is much wider than that of quantities, as evident from the distribution of (log) consumer types in Table 2. Overall, marginal utility $\theta(q)\nu'(q)$ decreases with q, although less fast than base marginal utility given that consumers' marginal willingness to pay increases with quantity.[33]

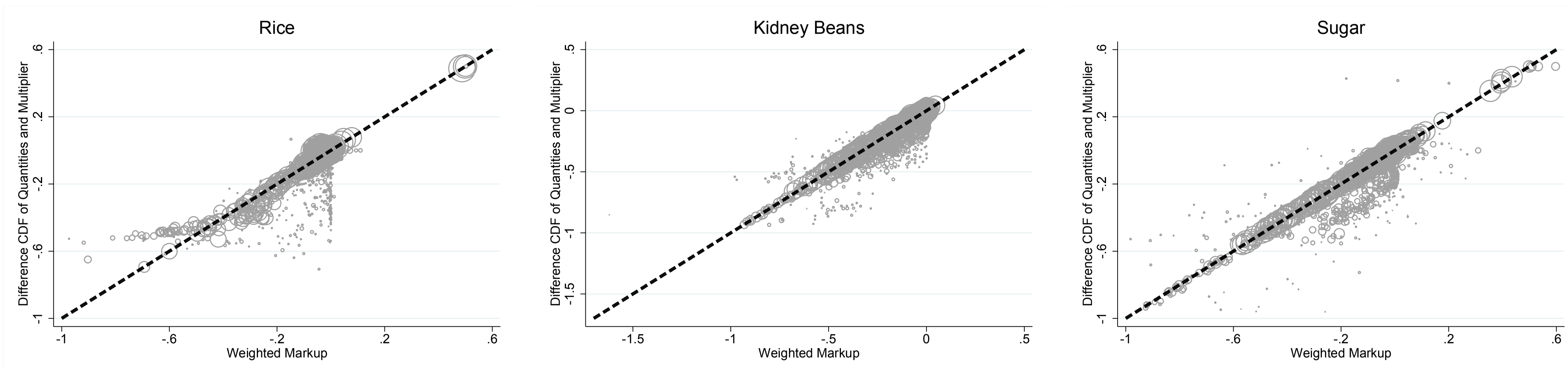
Figure 3: Estimated Base Marginal Utility (Top Panels) and Marginal Utility (Bottom Panels)



The large curvature of marginal utility that we estimate suggests the potential for rich distributional implications of nonlinear pricing. The reason is not only that consumers markedly differ in their tastes and so in their marginal willingness to pay, but also that different quantities are valued quite differently by consumers of any given taste. We explore these implications of nonlinear pricing in Section 5.3, where we link the scope for price discrimination, as captured by consumers' tastes and marginal utility, to the type of price discrimination that we infer sellers practice in our villages.

[33]The estimated reverse hazard rate $f(\theta)/F(\theta)$ and hazard rate $f(\theta)/[1 - F(\theta)]$ of the distribution of types in each village mostly decrease with θ for each good, which is a sufficient condition for (PS); see the Supplementary Appendix. As neither of these restrictions has been imposed in estimation, we interpret these findings as validating our estimates of the type distribution.

Figure 4: Model Fit Within and Across Villages



Model Fit. By a seller's first-order condition in (9), a key implication of our model is that the shape of the price schedule in a village is determined by the cumulative distribution function of consumers' marginal willingness to pay $F(\theta)$, which satisfies $F(\theta) = G(q)$ for $q = q(\theta)$, the associated support $[\underline{\theta}, \overline{\theta}]$ and probability density function $f(\theta)$, the multiplier $\gamma(\theta(q))$ on consumers' participation (or budget) constraints, and marginal cost $c'(Q)$. Although the distribution of marginal willingness to pay, the multiplier, and marginal cost are all unobserved, they are directly related to the observed distributions of quantities and unit prices by (17). Thus, one way to assess the fit of the model to the data is to determine the extent to which our estimates of $c'(Q)$, the auxiliary function $x(q)$, and $\gamma(\theta(q))$ satisfy the relationship between the observed distribution of quantities $G(q)$ and unit prices $p(q)$ implied by (17). To this purpose, for each good we plot in Figure 4 the estimated value of $G(q)-\gamma(\theta(q))$ on the y-axis against the estimated markup measure $1/[t_1 p(q)]-1/c'(Q)$ weighted by the auxiliary function $x(q)$, or *weighted markup* for brevity, on the x-axis. A point in any plot represents the fit of the model to the data for a particular purchased quantity in a village. Then, 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. Figure 4 shows that the model fits well the price and quantity data. For instance, the R^2 of a linear regression of $G(q)-\gamma(\theta(q))$ on the weighted markup is 0.927, 0.951, and 0.957 for rice, kidney beans, and sugar, respectively.

5.3 Distributional Implications of Nonlinear Pricing

Here we evaluate the degree of inefficiency of observed nonlinear pricing and assess its desirability by analyzing a counterfactual scenario in which sellers are prevented from discriminating and so price linearly.[34]

5.3.1 Distortions Associated with Price Discrimination

As discussed, our model is consistent with different degrees of market power among sellers. Sellers' market power not only can distort the allocation of a good relative to the first best, thereby reducing the gains from trade, but also can affect 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. The seller, however, would obtain all surplus. Alternatively, a seller 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 a larger surplus share accrues to consumers.

Here we examine the size of the distortions induced by sellers' market power, as implied by our estimates, through two measures: the percentage difference between the estimated multiplier $\gamma(\theta(q))$ and the cumulative distribution function of quantities purchased $G(q)$, since this difference would be zero under the first best by (9) and (17), and the fraction of households who purchase quantities larger than under the first best, namely, with $\gamma(\theta(q)) > G(q)$.[35] In the first five columns of Table 3, we report the average per-

[34] To reduce the impact of extreme observations, we exclude from this analysis of each good villages in which consumer types are estimated to be implausible large, namely, in which the first quartile of the distribution of households' marginal willingness to pay exceeds one million, winsorize the top 10% of the distribution of consumer types in each village, and focus on the resulting villages with at least two distinct quantities purchased. See the Supplementary Appendix for details.

[35] We interpret the first best as a scenario in which free entry in a market is possible, so sellers price at cost. We compute the percentage difference between x and x' as the ratio of $(x' - x)$ to $(x' + x)/2$.

centage deviation of $\gamma(\theta(q))$ from $G(q)$ in absolute value across villages ("Social Surplus Distortion") for selected percentiles in the distribution of consumer types in each village, namely, for households with types below the 5th percentile in the distribution of types in each village (first column), between the 5th and the 25th percentiles (second column), between the 25th and the 50th (third column), between the 50th and the 75th (fourth column), and above the 75th (fifth column). In the remaining five columns, we report the percentage of households across villages who consume *above* the first best ("Overconsumption") for the same percentiles in the distribution of types.

Table 3: Nonlinear Pricing vs. Perfectly Competitive Pricing by Percentile Ranges of Consumer Types

	Social Surplus Distortion					Overconsumption				
Good	5%	25%	50%	75%	100%	5%	25%	50%	75%	100%
Rice	136.7	123.3	62.6	39.6	17.2	1.4	2.9	6.1	29.6	87.3
Kidney Beans	168.4	120.5	51.6	22.6	9.1	0.0	2.2	9.4	12.5	69.6
Sugar	148.7	79.9	39.9	26.8	6.4	2.9	2.2	7.8	17.6	91.2

As apparent from the first five columns of Table 3, the distortions associated with nonlinear pricing are larger for households who consume low to intermediate quantities in each village, that is, up to the median consumer type. Households who consume larger quantities suffer much smaller distortions. As evident from the last five columns of the table, the consumption of households with low to intermediate types is also most compressed relative to the first best. Perhaps surprisingly, though, a small fraction of consumers with types below the median consume quantities *above* the first best. A much larger fraction of consumers with intermediate to large types consume above the first best, especially those with the greatest taste for a good who purchase the largest quantities. Hence, overall sellers practice an inefficient form of price discrimination, which, however, leads several households to overconsume rather than, as often suggested, underconsume.

5.3.2 Nonlinear versus Linear Pricing

It has been argued that the ability of sellers to price discriminate through quantity discounts hurts poor consumers. In particular, quantity discounts may limit the access of the poorest households to basic goods and services, as these households tend to purchase the smallest quantities and so face the highest unit prices (see Attanasio and Frayne (2006) for references). Based on our estimates, we can examine which households benefit more from the price discrimination we observe by comparing each household's consumer surplus and level of consumption under observed nonlinear pricing and under the counterfactual scenario that would emerge if sellers had market power but were constrained to price linearly—for instance, by regulation. Importantly, this exercise entails comparing not only equilibrium prices and quantities under the two pricing schemes but also the *size* of the market served under each: by Proposition 3, a seller who is prevented from discriminating may end up excluding some consumers under linear pricing. We find this to be the case for many villages in our sample.

Since this exercise requires considering quantities purchased outside of the observed range, we parameterize the estimated base marginal utility function in each village as a three-parameter HARA function,

$\nu'(q) = a[aq/(1-d)+b]^{d-1}$ with $a > 0$ and $aq/(1-d)+b > 0$.[36] To determine which households would participate under linear pricing, we also need an estimate of consumers' reservation utility. As discussed, though, reservation utility is identified only 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 under nonlinear pricing if the multiplier on this type's participation (or budget) constraint is estimated to be significantly different from zero, which implies that the relevant constraint binds. We then specify any higher type's reservation utility as equal to such type's estimated utility, if the multiplier on such type's constraint significantly differs from that of the next-lower type, and equal to the next-lower type's estimated utility otherwise.[37]

Based on these marginal utility and reservation utility schedules, in the first five columns of Table 4, we report the percentage of households across villages whose consumer surplus is higher under linear pricing than under nonlinear pricing for five groups: households with types below the 5th percentile in the distribution of consumer types in each village (first column), between the 5th and the 25th percentiles (second column), between the 25th and the 50th (third column), between the 50th and the 75th (fourth column), and above the 75th (fifth column). In the last five columns, we report the percentage of households across villages who consume more under linear pricing that under nonlinear pricing for the same percentiles in the distribution of consumer types.

Table 4: Linear Pricing (LP) vs. Nonlinear Pricing (NLP) by Percentile Ranges of Consumer Types

	Consumer Surplus under LP vs. NLP					Consumption under LP vs. NLP				
Good	5%	25%	50%	75%	100%	5%	25%	50%	75%	100%
Rice	79.6	88.2	81.1	87.8	96.4	46.9	51.5	46.3	65.8	89.1
Kidney Beans	30.1	23.7	26.6	23.2	55.3	18.7	3.7	3.6	3.6	49.1
Sugar	55.0	45.2	47.3	41.7	76.5	25.7	14.0	5.4	3.5	50.0

As the first five columns of the table show, except for rice, linear pricing leads to *lower* consumer surplus for most households in the first three quartiles of the distribution of consumer types in each village. On the contrary, households with the greatest taste for kidney beans and sugar tend to benefit from linear pricing. As the last five columns of the table show, except for rice, the overwhelming majority of consumers in the first three quartiles would also consume *less* under linear pricing. Results for rice are different since the estimated marginal utility for rice is smaller over the range of quantities that most households buy, as apparent by comparing the bottom panels of Figure 1 and the top panels of Figure 3. Intuitively, the lower $\nu'(q)$, the lower $A(q)$ since $\nu'(q) = a^d A(q)^{1-d}$ by the HARA form assumed—we estimate that $d < 1$ for each good—and so the higher $A(q)^{-1}$. In turn, a higher average $A(q)^{-1}$ implies a higher price elasticity of

[36] Specifically, we estimate the parameter d from a pooled regression of estimated base marginal utility on quantity with $a = 1-d$ to limit parameter proliferation. We estimate, instead, a and b from analogous village-level regressions and focus on villages for which the corresponding adjusted R^2 is at least 0.75. For this exercise, we interpret $c'(Q)$ as the marginal cost of a cost function with zero fixed costs and constant marginal cost.

[37] If the multiplier on type θ_i's constraint significantly differs from that on type θ_{i-1}'s constraint, there exists at least one type between θ_{i-1} and θ_i whose constraint binds. Because of the sparseness of the data, we interpret such type to be θ_i. In the Supplementary Appendix, we consider the case in which the reservation utility is the maximal possible for each type, that is, $\overline{u}(\theta_i)=u(\theta_i)$, and obtain similar results.

aggregate demand in absolute value, since this elasticity can be expressed as

$$|\varepsilon_{QP}| = \frac{E_\theta[A(q_m(\theta))^{-1}]}{E_\theta[q_m(\theta)]} = \frac{1}{1-d} + \frac{b}{aE_\theta[q_m(\theta)]}$$

and thus increases with $E_\theta[A(q_m(\theta))^{-1}]$ and, given that (the median) b is estimated to be small, with d. In fact, the median price elasticity of aggregate demand under linear pricing is largest for rice whereas the median marginal cost is smallest for rice. Hence, unlike for kidney beans and sugar, in the case of rice, sellers do not have an incentive to charge high prices in the hope of attracting consumers with large valuations who are willing to pay more. Indeed, both the mean and the median linear price across villages are lowest for rice. As a result, households are largely better off under linear pricing.

A key reason why consumer surplus is higher under nonlinear pricing for kidney beans and sugar is the higher degree of market participation associated with nonlinear pricing. We measure this effect by the percentage of households across villages who would not participate in the market under linear pricing. The percentages of excluded households in the percentile ranges of Table 4 are, respectively, 20.4%, 11.8%, 18.3%, 10.7%, and 0.9% for rice; 69.9%, 74.8%, 70.8%, 71.4%, and 18.0% for kidney beans; and 45.0%, 54.8%, 51.5%, 52.3%, and 9.6% for sugar. Thus, a large fraction of households purchasing kidney beans and sugar in the first three quartiles of the distribution of consumer types in each village would be excluded under linear pricing, whereas nearly all households participate under observed nonlinear pricing, as discussed in Section 2. The logic behind this result is simple. In the case of kidney beans and sugar, the relatively low price elasticity of aggregate demand in absolute value implies that high linear prices are optimal for sellers, even if they lead consumers with low taste parameters to opt out of the market. Given the much higher marginal willingness to pay of high types relative to low types reported in Table 2, sellers more than make up for excluding low types by charging high prices to the remaining ones.

5.4 The Effect of Income Transfers

In this subsection, we show that our model can account for a substantial fraction of the observed dispersion in unit prices across quantities within a village and across villages, as well as for the shift in the schedule of unit prices induced by Progresa, which we have documented in Table 1. Specifically, we first show that a reduced form of our model from a Taylor expansion of the seller's first-order condition in (9), in which the slope of unit price schedules depends on the hazard rate of the distribution of quantities purchased in each village, explains a large fraction of the variation in unit prices within and across villages. We then show that once the dependence of unit prices on these hazard rates is explicitly accounted for, the effect of the program on unit price schedules is no longer significant. Hence, our model is sufficient to account for the impact of the program.

Recall from Proposition 4 that 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 $T(q)$ to some 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 unit prices of agricultural commodities. As mentioned, however, Hoddinott 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 prices of basic staples. Unlike these studies that focus only on the impact of transfers on average unit prices, in Section 2 we have examined the impact of Progresa on their entire schedule. In Table 1, we have documented that Progresa has had a significant effect on unit prices in that it has lead to an increase in the magnitude of quantity discounts but that this effect cannot be detected without taking the nonlinearity of unit prices into account, as consistent with our model. Specifically, our model implies that the effect of an income transfer on unit prices, $p(q)=T(q)/q$, is ambiguous, since both households' expenditure, $T(q)$, and consumption, q, typically increase. Although average unit prices may not increase after a transfer, the price schedule can substantially change as we proved in Corollary 1, leading overall to a greater intensity of price discrimination, which we observe in our data. Here we show that our model can explain such a change in unit prices.

Transfers and Prices. We examine 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 in (9),

$$\log[p(q)]\approx\beta_0+\beta_1\log(q)+\beta_2\left[\frac{1-G(q)}{g(q)}\right]+\beta_3\log(q)\left[\frac{1-G(q)}{g(q)}\right]+\beta_4\log(q)^2+\beta_5\left[\frac{1-G(q)}{g(q)}\right]^2, \quad (19)$$

derived at the end of Appendix A. In this expansion, the multiplier $\gamma(\theta(q))$ is interpreted 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. This latter term captures the importance of the shape of the distribution of consumers' marginal willingness to pay for unit prices. Intuitively, according to our model, unit prices are related to the *distribution* of consumer preferences, in particular to its inverse hazard rate, which is apparent by rewriting the right side of (9) as the product of $1/\theta$ and $[\gamma(\theta(q))-1+1-F(\theta)]/f(\theta)$. By the one-to-one relationship between consumer tastes and demand, unit prices are then related to the hazard rate of the distribution of quantities purchased in a village, as (19) shows.

In Table 1, we documented a significant shift in the price schedules of the three goods of interest after the Progresa transfer. In Table 5, we assess the extent to which our model can account for the observed nonlinearity of prices as well as for the change in the unit price schedules resulting from Progresa. To this end, we first estimate (19) and then a version of it augmented to incorporate the impact of the Progresa transfer through a "Treatment" dummy and the interaction between this dummy and log quantity. We stress that the inverse hazard rate of quantities $[1-G(q)]/g(q)$ in both regressions is computed for each village. This approach thus allows for heterogeneous impacts of the program across villages and could be interpreted as a *mediation analysis* of the effect of Progresa on price schedules, since the program has affected the hazard rate of the distribution of quantities in treated villages. For instance, relative to control villages, the inverse hazard rate $[1-G(q)]/g(q)$ in treated villages on average is 18.6% higher for rice, 10.8% higher for kidney

Table 5: Impact of Cash Transfers on Prices (98% Trimming)

	Rice Unit Values		Kidney Beans Unit Values		Sugar Unit Values	
	1	2	1	2	1	2
Intercept	1.877***	1.876***	2.454***	2.456***	1.792***	1.787***
	(0.006)	(0.008)	(0.008)	(0.011)	(0.004)	(0.006)
Treatment		0.001		-0.003		0.006
		(0.009)		(0.012)		(0.006)
$\log(q)$	-0.140***	-0.135***	-0.222***	-0.216***	-0.197***	-0.193***
	(0.009)	(0.013)	(0.015)	(0.018)	(0.010)	(0.013)
$\frac{1-G(q)}{g(q)}$	-0.002	-0.002	-0.004***	-0.004***	0.000	0.000
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
$\log(q) \times \frac{1-G(q)}{g(q)}$	-0.005***	-0.005***	-0.005***	-0.005***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\log(q) \times$Treatment		-0.006		-0.007		-0.005
		(0.012)		(0.012)		(0.011)
$\log(q)^2$	0.118***	0.118***	0.082***	0.082***	0.107***	0.107***
	(0.009)	(0.009)	(0.012)	(0.012)	(0.009)	(0.009)
$\left[\frac{1-G(q)}{g(q)}\right]^2$	-0.000	-0.000	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.419	0.420	0.278	0.278	0.330	0.330
Observations	69,543	69,543	93,375	93,375	103,930	103,930

Note: * for $p < 0.10$, ** for $p < 0.05$ and *** for $p < 0.01$. Clustered standard errors. Wave fixed effects included.

beans, and 17.6% higher for sugar for small quantities, namely, those in the top 25% of the distribution of unit prices.[38]

In columns 1, we report the estimates of (19). The effect of the interaction of $\log(q)$ with the inverse hazard rate $[1 - G(q)]/g(q)$ is significant for each of the three goods at the 1% level; for kidney beans, the effect of the inverse hazard rate is also significant. The effect of the quadratic inverse hazard rate is not significant for any good, whereas that of the quadratic term in $\log(q)$ is significant for all. Overall, this specification accounts for a large fraction of the dispersion in unit prices within and between villages.

In columns 2, we report the estimates of an augmented version of equation (19) 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 does not significantly affect the dependence of unit prices on $\log(q)$ either. That is, the point estimates of the coefficient on $\log(q)$ interacted with the "Treatment" dummy are greatly reduced in absolute value relative to the estimates reported in columns 3 of Table 1 and are no longer significant for any good.

On the contrary, the interaction between $\log(q)$ and $[1 - G(q)]/g(q)$ is estimated to be significant for each good. Hence, this statistic is sufficient to account for the effect of the program on unit prices in a precise sense: conditional on the interaction between $\log(q)$ and $[1 - G(q)]/g(q)$, that between $\log(q)$ and "Treatment" no longer significantly affects price schedules. Moreover, comparing the estimated coefficients

[38] Standard errors are computed by bootstrap at the village level to account for the fact that hazard rates are estimated for each locality and wave. We could allow for village fixed effects to capture the unobserved variability in marginal costs across villages—since we have several waves of data, both wave fixed effects and village fixed effects are identified. The results we obtain allowing for village fixed effects are very similar to those in Table 5. See the Supplementary Appendix, where we also report analogous results for alternative stratification and clustering schemes.

on the interaction between $\log(q)$ and the "Treatment" dummy in columns 3 in Table 1 and those on the interaction between $\log(q)$ and $[1 - G(q)]/g(q)$ in columns 2 in Table 5 reveals that a significant portion of the change in prices induced by the program is accounted for by the change in the distribution of quantities purchased in each village, in particular by a change in their curvature, as captured by the inverse hazard $[1-G(q)]/g(q)$. These results thus indicate that our model is capable of explaining the shift in price schedules documented in Table 1.

These findings support the key implication of our model that unit prices vary with quantity and that the relationship between unit prices and quantity is affected by the distribution of consumer tastes, and so quantities purchased, across consumers. Based on the results in columns 3 of Table 1, the program did not lead to a significant change in average unit prices between control and treated villages. Unit prices, though, have changed substantially and differentially for consumers of small and large quantities. For instance, unit prices in the top 25% of the unit price distribution across treated villages, paid by the households who purchase small quantities, on average are 17.6% *higher* for rice, 38.8% higher for kidney beans, and 52.5% higher for sugar than across control villages. On the contrary, unit prices in the bottom 25% of the unit price distribution across treated villages, paid by the households who purchase large quantities, on average are 22.6% *lower* for rice, 22.1% lower for kidney beans, and 14.0% lower for sugar than across control villages.[39] (All these differences between treated and control villages are significant at the 1% level.) Relative to control villages, the average village-level standard deviation of unit prices across treated villages is 77.8% higher for rice, 129.7% higher for kidney beans, and 124.8% higher for sugar.

Since all households in the villages receiving the Progresa transfer have been affected by these price changes and in varying degrees depending on the quantities they purchase, the transfer has had an indirect and non-uniform price effect on non-eligible households.[40] In light of the increase in the intensity of discrimination that we document, then, the transfer may have had a more limited beneficial impact overall than commonly inferred.

6 Conclusion

We propose a model of nonlinear pricing in which consumers differ in their taste 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 advancement in the literature, since it makes the model relevant for several contexts of practical relevance in developing countries. In the settings we consider, the distributional effects of nonlinear pricing across consumers are fundamentally different from those arising from standard models of nonlinear pricing, in which consumers are assumed to be unconstrained in their purchase decisions and outside options are presumed to

[39]These changes have been even more pronounced for purchasers of the smallest and largest quantities. Qualitatively similar changes between treated and control villages can also be detected for log unit prices in terms of both their overall average and the average of various percentiles of their within-village distributions across villages. We have also estimated quantile treatment effects of the program on log unit prices and found significant changes between treated and control villages in line with the patterns discussed here. Results are similar when real expenditures and quantities are trimmed at the top 1% or 5% rather than at the top 2% as in Table 5. See the Supplementary Appendix for details.

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

be identical across consumers. In particular, in the model we propose, quantity discounts for *large* volumes can be associated with consumption above the first best at *low* volumes.

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

Crucially, we show that by increasing consumers' ability to pay, cash transfers provide sellers with the incentive to extract more surplus from consumers through nonlinear pricing. As a result, cash transfers can lead to an increase in the intensity of price discrimination, as we document in the case of Progresa. A few studies have analyzed the effect of transfers on the price of commodities, and the consensus so far seems to be that Progresa did not have appreciable effects on local unit prices. We estimate, instead, that the cash transfers implemented by Progresa have had a significant impact on unit prices in our villages by inducing a shift in price schedules. We also show that our model can explain not only a large fraction of the dispersion in unit prices within and across villages but also the observed shifts in price schedules. Specifically, although the program has not affected unit prices on average, we document that once the dependence of unit prices on quantity is taken into account, the price effect of the program is substantial. In particular, the program is associated with a significant increase in the degree of price discrimination, which our model accounts for.

Our paper is one of the first to uncover important shifts in price schedules in villages included in the Progresa evaluation sample. This result is all the more relevant since cash transfers have become an increasingly popular poverty alleviation measure in Latin America and many other developing countries. Our estimation results thus suggest the importance of accounting not only for heterogeneity in consumers' preferences, constraints, and consumption opportunities but also for the nonlinearity of prices when assessing the impact of cash transfers.

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?," Unpublished Manuscript.

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

BONTEMPS, C., and D. MARTIMORT (2014): Identification and Estimation of Incentive Contracts under

Asymmetric Information: An Application to the French Water Sector," Unpublished Manuscript.
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.
CUNHA, J., G. DE GIORGI, and S. JAYACHANDRAN (2017): "The Price Effects of Cash Versus In-Kind Transfers," Unpublished Manuscript.
DEATON, A., (1989): "Rice Prices and Income Distribution in Thailand: A Non-parametric Analysis," *Economic Journal*, 99(395), 1-37.
DHARMADHIKARI, S.W., and K. JOAG-DEV (1983): "Mean, Median, Mode III," *Statistica Neerlandica*, 37(4), 165–168.
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," Unpublished Manuscript.
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.
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.
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," Unpublished Manuscript.
SEGAL, I., and S. TADELIS (2005): "Lectures in Contract Theory," Unpublished Manuscript.
STOLE, L.A. (2007): "Price Discrimination and Competition," in *Handbook of Industrial Organization*, Vol. 3, ed. by M. Armstrong and R. Porter, 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," Unpublished Manuscript.

A Model: Omitted Proofs and Details

Proof of Proposition 1: Before proving Proposition 1, we first derive the simple BC problem in (6) and then establish that the first-order and complementary slackness conditions of the simple BC problem are necessary and sufficient to characterize an optimal menu. The proof of these results requires that assumptions analogous to (PS), (H), and (FP) in the IR model hold in the BC model. We have discussed assumptions (BCH) and (FP) in the main text, so here we discuss only assumption (PS). 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 $\Phi \in [0, 1]$. In the IR model, sufficient conditions for (PS) 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, the monotonicity condition for $q(\theta)$ for incentive compatibility is easier to satisfy. The second and third inequalities in (20) simply amount to a strengthening of the usual monotone hazard rate condition of nonlinear pricing models.

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 constraint expressed as in (21). Next, note that by adding and subtracting $\int_{\underline{\theta}}^{\overline{\theta}} u(\underline{\theta})f(\theta)d\theta$, we obtain

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

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

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

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

Similarly,

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

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

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

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

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

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 an argument similar to the one in the proof of Result 1 in the Supplementary Appendix, it is possible to show that $\Phi(\overline{\theta})=1$. Then, 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 it is also a solution to the BC problem. For notational simplicity, in the following we suppress the subscript *IR* from $u_{IR}(\theta)$, $q_{IR}(\theta)$, $t_{IR}(\theta)$, $\overline{u}_{IR}(\theta)$, and $\overline{q}_{IR}(\theta)$. 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

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

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

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

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 $u(\theta)\geq\overline{u}$; see Result 4 in the Supplementary Appendix for the uniqueness of the optimal allocation in the IR problem. Note that for $\Phi(\theta)$ to be a legitimate cumulative multiplier, it must be nonnegative and weakly increasing with θ. Let $\Phi(\theta)=\gamma(\theta)$ be the cumulative multiplier in the BC problem. Clearly, $\Phi(\theta)=\gamma(\theta)$ is a legitimate cumulative multiplier and is such that 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 $u(\theta) \geq \overline{u}$. We first claim that at the IR allocation, the complementary slackness condition of the BC problem, (29), holds and the IR allocation satisfies $t(\theta) \leq I(\theta, q(\theta))$. To this purpose, recall that $I(\theta, q(\theta)) \geq v(\theta, q(\theta)) - \overline{u}(\theta)$, $I(\theta, \overline{q}(\theta)) = v(\theta, \overline{q}(\theta)) - \overline{u}(\theta)$, 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 $u(\theta) = v(\theta, q(\theta)) - t(\theta) = \overline{u}(\theta)$ so that $t(\theta) = v(\theta, q(\theta)) - \overline{u}(\theta)$. Since, by assumption, $v(\theta, q(\theta)) - \overline{u}(\theta) = I(\theta, q(\theta))$ for types whose IR constraints bind, it follows that $t(\theta) = I(\theta, q(\theta))$ for such types. When, instead, the IR constraints do not bind so that $d\gamma(\theta) = d\Phi(\theta) = 0$, then $u(\theta) = v(\theta, q(\theta)) - t(\theta) \geq \overline{u}(\theta)$ so that $t(\theta) \leq v(\theta, q(\theta)) - \overline{u}(\theta)$. 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 $u(\theta) \geq \overline{u}$ by (IR) and the fact that $\overline{u}(\underline{\theta}) \geq \overline{u}$ by assumption.

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. Recall that $I_q(\theta, q(\theta))$ equals $v_q(\theta, q(\theta))$ when the IR constraints bind by assumption. Thus, either $\phi(\theta) = 0$ or, if not, $I_q(\theta, q(\theta)) = v_q(\theta, q(\theta))$ for each θ. Hence, the second term on the right side of (28) equals zero for each θ, and so the first-order conditions of the BC problem in (28) are identical to those of the IR problem in (26).

Step 3: Verify IR and BC Problems Imply Same Utility. 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, consider a type θ' whose IR constraint binds. Then, in the IR problem for any type θ higher than θ', we have

$$u(\theta) = \overline{u}(\theta') + \int_{\theta'}^{\theta} v_\theta(x, q(x))dx = v(\theta', \overline{q}(\theta')) - I(\theta', \overline{q}(\theta')) + \int_{\theta'}^{\theta} v_\theta(x, q(x))dx, \tag{30}$$

since $u'(\theta) = v_\theta(\theta, q(\theta))$ by local incentive compatibility, and $u(\theta') = \overline{u}(\theta') = v(\theta', \overline{q}(\theta')) - I(\theta', \overline{q}(\theta'))$ by assumption and construction of θ'. The utility in (30) 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.

Thus, the solutions to the IR and BC problems are the same. This establishes the desired result. □

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, based only 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 allocation 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 in submarket w 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) are satisfied with $I(\theta, q, w) = Y(w) - \underline{z}(\theta, q)$. Our next result shows how this menu varies across submarkets. For this, let

$$t(\theta, w_H) = t(\theta, w_L) + Y(w_H) - Y(w_L),\ q(\theta, w_H) = q(\theta, w_L), \text{ and } \Phi(\theta, w_H) = \Phi(\theta, w_L). \tag{31}$$

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 (31) 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 is positive for type (θ, w_H) in submarket w_H if, and only if, it is positive 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,

$$z(\theta, w_L) = \underline{z}(\theta, q(\theta, w_L)) = Y(w_L) - t(\theta, w_L). \tag{32}$$

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 (31) satisfy 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, the first-order conditions in the two submarkets are identical under (31). 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

$$t(\theta, w_L) = I(\theta, q(\theta, w_L), w_L) \equiv Y(w_L) - \underline{z}(\theta, q(\theta, w_L)). \tag{33}$$

But then the multiplier $\phi(\theta, w_H)$ in submarket w_H for this same type θ is also positive under (31), 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 (31) and the second equality from (33). Hence, the conjectured solution satisfies the first-order conditions and complementary slackness condition for submarket w_H. So, by Result 1, the conjectured allocation 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 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

$$I(\theta, q, \omega(\theta)) = Y(\omega(\theta)) - \underline{z}(\theta, q) \tag{34}$$

for any q. Under (34), the analogues of Result 1 and Proposition 1 apply. □

Proof of Proposition 2: Recall that $T'(q(\theta)) = \theta\nu'(q(\theta)) > 0$ by local incentive compatibility, and note that $A(q) = -\nu''(q)/\nu'(q) > 0$ since $\nu'(\cdot) > 0$ and $\nu''(\cdot) < 0$ by assumption. Differentiating $T'(q) = \theta(q)\nu'(q)$ yields

$$T''(q) = \theta'(q)\nu'(q) + \theta(q)\nu''(q) = \theta(q)\nu'(q)\left[\frac{\theta'(q)}{\theta(q)} + \frac{\nu''(q)}{\nu'(q)}\right] = T'(q)\left[\frac{1}{\theta(q)q'(\theta)} - A(q)\right]. \tag{35}$$

By using $T'(q) = \theta(q)\nu'(q)$, the first-order condition in (9) can be expressed as $\{\theta - [\gamma(\theta) - F(\theta)]/f(\theta)\}\nu'(q(\theta)) - c = 0$.

Applying the implicit function theorem to this latter 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 $\theta > [\gamma(\theta) - F(\theta)]/f(\theta)$ by the seller's first-order condition since $\nu'(\cdot) > 0$. Hence, the denominator of $q'(\theta)$ is positive. Using (35) and the fact that $T'(q), A(q) > 0$, we can equivalently express $T''(q) \leq 0$ as

$$T'(q(\theta))A(q(\theta))\left\{\frac{\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}}{\theta\frac{\partial}{\partial\theta}\left[\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}\right]} - 1\right\} \leq 0 \Leftrightarrow \frac{\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}}{\theta\frac{\partial}{\partial\theta}\left[\theta - \frac{\gamma(\theta)-F(\theta)}{f(\theta)}\right]} \leq 1. \tag{36}$$

Consider first any interval of consumer types whose IR constraints bind. By construction, any such type θ consumes $\overline{q} = \overline{q}(\theta)$ and achieves utility $\overline{u}(\theta)$. Assumption (H) implies that $\overline{u}'(\theta) = \nu(\overline{q}(\theta))$, which in turn yields that $\overline{q}'(\theta) = \overline{u}''(\theta)/\nu'(\overline{q}(\theta))$. Then, by (35)

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

Since $\nu'(\cdot) > 0$, it follows that $T''(\overline{q}(\theta)) \leq 0$ if, and only if, $\nu'(\overline{q}(\theta)) \leq \theta A(\overline{q}(\theta))\overline{u}''(\theta)$, 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 such types. When $\gamma = 1$, it follows that

$$q'(\theta) = \frac{\frac{\partial}{\partial\theta}\left[\theta - \frac{1-F(\theta)}{f(\theta)}\right]}{\left[\theta - \frac{1-F(\theta)}{f(\theta)}\right]A(q(\theta))} \geq \frac{1}{\theta A(q(\theta))} = \frac{\nu'(q(\theta))}{-\theta\nu''(q(\theta))}, \tag{37}$$

where the inequality follows from the assumption that $[1 - F(\theta)]/f(\theta)$ decreases with θ and the fact that $\theta > [1 - F(\theta)]/f(\theta) \geq 0$. Condition (37) implies that $1/q'(\theta) \leq -\theta\nu''(q(\theta))/\nu'(q(\theta))$, which combined with (35) yields that

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

When, instead, $\gamma \in [0, 1)$, the last inequality in (36) becomes

$$\frac{\theta f(\theta) - \gamma + F(\theta)}{\theta f(\theta)} \leq \frac{\partial}{\partial\theta}\left[\theta - \frac{\gamma}{f(\theta)} + \frac{F(\theta)}{f(\theta)}\right] \Leftrightarrow \theta f^2(\theta) \geq -[\gamma - F(\theta)]f(\theta) - [\gamma - F(\theta)]\theta f'(\theta). \tag{38}$$

We prove that (38) 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 (38) is

$$f^2(\theta) \geq -[\gamma - F(\theta)]f'(\theta). \tag{39}$$

If $f'(\theta) \geq 0$, then it is immediate that (39) 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 $f'(\theta) < 0$ and $\gamma < 1$, it follows that (39) and so (38) are satisfied.

Case 2: $\gamma < F(\theta)$. Note that we can rewrite (38) as

$$\theta f(\theta) \geq F(\theta) - \gamma + [F(\theta) - \gamma]\theta\frac{f'(\theta)}{f(\theta)}. \tag{40}$$

When $f'(\theta) \leq 0$, a sufficient condition for (40) is $F(\theta) \leq \theta f(\theta)$, which is assumed. Thus, (38) holds. When, instead, $f'(\theta) > 0$, a sufficient condition for (40) is $\theta f(\theta) \geq F(\theta) + F(\theta)\theta f'(\theta)/f(\theta)$ 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 holds by assumption. Hence, (38) is satisfied. □

Proof of Proposition 3: We first establish the claim under 1) and then the claim under 2).

1) We divide the proof of this claim 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 $q_m(\theta) \geq q(\theta)$, then the utility of a consumer of type θ is higher under linear pricing 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

$$u(\theta) = \theta\nu(q(\theta)) - T(q(\theta)) > u_m(\theta) = \theta\nu(q_m(\theta)) - \theta\nu'(q_m(\theta))q_m(\theta), \tag{41}$$

where in (41) 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

$$\theta\nu(q_m(\theta)) - \theta\nu'(q_m(\theta))q_m(\theta) \geq \theta\nu(q(\theta)) - \theta\nu'(q_m(\theta))q(\theta), \tag{42}$$

since 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. Note that (41) and (42) imply that $\theta\nu(q(\theta)) - T(q(\theta)) > \theta\nu(q(\theta)) - \theta\nu'(q_m(\theta))q(\theta)$, that is,

$$\theta\nu'(q_m(\theta)) > T(q(\theta))/q(\theta) = p(q(\theta)). \tag{43}$$

Now, by the assumption of the case, $p'(q) = [T'(q) - T(q)/q]/q \leq 0$ at $q = q(\theta)$ or, equivalently,

$$T'(q(\theta)) \leq T(q(\theta))/q(\theta). \tag{44}$$

This inequality, together with (43) and the incentive compatibility condition $T'(q(\theta)) = \theta\nu'(q(\theta))$, implies

$$\theta\nu'(q(\theta)) = T'(q(\theta)) \leq T(q(\theta))/q(\theta) < \theta\nu'(q_m(\theta)), \tag{45}$$

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 weakly decreasing. Hence, $u_m(\theta) \geq u(\theta)$.

Case b). We now show that if the price schedule exhibits quantity discounts in that $T''(q) \leq 0$ at 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 so, then we contradict the assumption that all types participate under linear pricing by showing that there exists a type $\theta_2 > \hat{\theta}$ whose participation constraint binds under nonlinear pricing, that is, $u(\theta_2) = \overline{u}(\theta_2)$, 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.

Consider then a consumer of type $\theta_2 > \hat{\theta}$ with $u(\theta_2) = \overline{u}(\theta_2)$. Note that such a consumer must exist if $\gamma(\hat{\theta}) < 1$. To reach the desired contradiction, rewrite $u(\hat{\theta}) > u_m(\hat{\theta})$ as

$$u(\theta_2) - [u(\theta_2) - u(\hat{\theta})] > u_m(\theta_2) - [u_m(\theta_2) - u_m(\hat{\theta})], \tag{46}$$

which can equivalently be expressed as

$$u(\theta_2) - \int_{\hat{\theta}}^{\theta_2} u'(x)dx > u_m(\theta_2) - \int_{\hat{\theta}}^{\theta_2} u'_m(x)dx. \tag{47}$$

Condition (47), in turn, is equivalent to

$$\overline{u}(\theta_2) - u_m(\theta_2) > \int_{\hat{\theta}}^{\theta_2} [\nu(q(x)) - \nu(q_m(x))]\, dx \tag{48}$$

by using $u(\theta_2) = \overline{u}(\theta_2)$ since the IR constraint of type θ_2 binds under nonlinear pricing by construction, by exploiting incentive compatibility under nonlinear pricing and consumer optimality under 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 (48) is positive, which establishes the desired contradiction. To see that the right side of (48) is positive, note first that for all $\theta \geq \hat{\theta}$,

$$p_m = \theta\nu'(q_m(\theta)) = \hat{\theta}\nu'(q_m(\hat{\theta})) \geq \hat{\theta}\nu'(q(\hat{\theta})) = T'(q(\hat{\theta})) \geq T'(q(\theta)) = \theta\nu'(q(\theta)), \tag{49}$$

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 weakly decreasing, the third and fourth equalities follow by 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 and $q(\theta) \geq q(\hat{\theta})$. Hence, (49) implies that $\theta\nu'(q_m(\theta)) \geq \theta\nu'(q(\theta))$ 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 weakly decreasing. But since $q(\theta) \geq q_m(\theta)$ for all $\theta \geq \hat{\theta}$ and $\nu(\cdot)$ is increasing, the right side of (48) is positive, which implies that $\overline{u}(\theta_2) > u_m(\theta_2)$. Then, θ_2 does not participate under linear pricing, which is a contradiction.

2) Consider now the proof of the claim under 2). 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 type under nonlinear pricing and these types participate; see Lemma 2 in Jullien (2000), which states that under (H), (FP) holds if $s(\theta, \overline{q}(\theta)) \geq \overline{u}(\theta)$ for all types. Let $\hat{\theta}$ be one such type. To establish the desired claim, we need to show that there exists a subinterval of types in $[\hat{\theta}, \theta'']$ who do not participate under linear pricing although, as just argued, they participate under nonlinear pricing. To this purpose, suppose, by way of contradiction, that all consumer types $[\hat{\theta}, \theta'']$ participate under linear pricing. We prove that if this is the case, then the seller makes negative profits under linear pricing. First, 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

$$\begin{aligned} u_m(\theta) &= u_m(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} u'_m(x)dx = u_m(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} \nu(q_m(x))dx \\ &\geq \overline{u}(\theta) = \overline{u}(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} \overline{u}'(x)dx = \overline{u}(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} \nu(\overline{q}(x))dx, \end{aligned} \tag{50}$$

where the second equality in (50) 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 assumption (H) in that $\overline{u}'(\theta) = \nu(\overline{q}(\theta))$. Since $u_m(\hat{\theta}) = \overline{u}(\hat{\theta})$ by assumption, (50) implies

$$\int_{\hat{\theta}}^{\theta} \nu(q_m(x))dx \geq \int_{\hat{\theta}}^{\theta} \nu(\overline{q}(x))dx. \tag{51}$$

With $\nu(\cdot)$ positive and increasing, (51) 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 decreasing for $\theta \in [\theta_3, \theta_4]$ gives

$$p_m = \theta\nu'(q_m(\theta)) < \theta\nu'(q^{FB}(\theta)) = c \tag{52}$$

for $\theta \in [\theta_3, \theta_4]$, where the first equality follows from the consumer's 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 (52) contradicts seller optimality: the seller can always raise p_m and earn at least zero profits. □

Proof of Proposition 4: Recall that $\{\overline{q}_{IR}(\theta)\}$ denotes the reservation quantity profile in the IR model and $\{t(\theta), q(\theta)\}$ denotes the optimal menu before the transfer is introduced with associated reservation utilities $\{\overline{u}(\theta)\}$ when consumers spend all of their budgets on the seller's good to purchase the incentive compatible quantities $\{\overline{q}(\theta)\}$. Denote by

$\{t_\tau(\theta), q_\tau(\theta)\}$ the optimal menu after the transfer is introduced with associated reservation utilities $\{\overline{u}_\tau(\theta)\}$ when consumers spend all of their budgets on the seller's good to purchase the incentive compatible quantities $\{\overline{q}_\tau(\theta)\}$. Before the transfer, for any consumer type whose BC constraint binds, the budget for the seller's good is $I(\theta, \overline{q}(\theta)) = Y - \underline{z}(\theta, \overline{q}(\theta))$ and utility is $\overline{u}(\theta) = \theta\nu(\overline{q}(\theta)) - Y + \underline{z}(\theta, \overline{q}(\theta))$ with

$$\overline{u}'(\theta) = \nu(\overline{q}(\theta)) + \theta\nu'(\overline{q}(\theta))\overline{q}'(\theta) + \underline{z}_\theta(\theta, \overline{q}(\theta)) + \underline{z}_q(\theta, \overline{q}(\theta))\overline{q}'(\theta).$$

For any consumer whose BC constraint bind, $\bar{t}(\theta) = I(\theta, \overline{q}(\theta))$ and so

$$\bar{t}'(\theta) = I_\theta(\theta, \overline{q}(\theta)) + I_q(\theta, \overline{q}(\theta))\overline{q}'(\theta) = -\underline{z}_\theta(\theta, \overline{q}(\theta)) - \underline{z}_q(\theta, \overline{q}(\theta))\overline{q}'(\theta). \tag{53}$$

After the transfer, for any consumer type whose BC constraint binds, the budget for the seller's good is

$$I(\theta, \overline{q}_\tau(\theta), \tau) = Y - \underline{z}(\theta, \overline{q}_\tau(\theta)) + \tau(\theta) = I(\theta, \overline{q}_\tau(\theta)) + \tau(\theta) \tag{54}$$

and utility is $\overline{u}_\tau(\theta) = \theta\nu(\overline{q}_\tau(\theta)) - Y + \underline{z}(\theta, \overline{q}_\tau(\theta)) - \tau(\theta)$ with

$$\overline{u}_\tau'(\theta) = \nu(\overline{q}_\tau(\theta)) + \theta\nu'(\overline{q}_\tau(\theta))\overline{q}_\tau'(\theta) + \underline{z}_\theta(\theta, \overline{q}_\tau(\theta)) + \underline{z}_q(\theta, \overline{q}_\tau(\theta))\overline{q}_\tau'(\theta) - \tau'(\theta).$$

For any consumer whose BC constraint binds, $\bar{t}_\tau(\theta) = I(\theta, \overline{q}_\tau(\theta), \tau)$. Using this condition and (54) yields that

$$\bar{t}_\tau'(\theta) = I_\theta(\theta, \overline{q}_\tau(\theta)) + I_q(\theta, \overline{q}_\tau(\theta))\overline{q}_\tau'(\theta) + \tau'(\theta) = -\underline{z}_\theta(\theta, \overline{q}_\tau(\theta)) - \underline{z}_q(\theta, \overline{q}_\tau(\theta))\overline{q}_\tau'(\theta) + \tau'(\theta). \tag{55}$$

By Proposition 1, $\overline{q}_{IR}(\theta) = \overline{q}(\theta)$ for types with binding BC constraints in the BC model. Then, before the transfer, the condition $\theta\nu'(\overline{q}(\theta))\overline{q}'(\theta) = \bar{t}'(\theta)$ by (BCH), (53), the requirement of Proposition 1 that $\theta\nu'(\overline{q}_{IR}(\theta)) = I_q(\theta, \overline{q}_{IR}(\theta))$, and the result that $\overline{q}_{IR}(\theta) = \overline{q}(\theta)$ imply that $I_\theta(\theta, \overline{q}(\theta)) = 0$. After the transfer, similarly, it must be that $I_\theta(\theta, \overline{q}_\tau(\theta)) + \tau'(\theta) = 0$ and so $I_\theta(\theta, \overline{q}_\tau(\theta)) = -\tau'(\theta) > 0$ given that $\tau'(\theta) < 0$. Since $I_\theta(\theta, \overline{q}_\tau(\theta)) > I_\theta(\theta, \overline{q}(\theta))$, it follows that $\overline{q}_\tau(\theta) > \overline{q}(\theta)$.

We now argue that there exists an interval of types with $q_\tau(\theta) \geq q(\theta)$ and $T_\tau'(q_\tau(\theta)) \leq T'(q(\theta))$. By assumption, there exists a consumer type θ' whose budget constraint binds before and after the transfer so that $q_\tau(\theta') = \overline{q}_\tau(\theta') > \overline{q}(\theta') = q(\theta')$ by the argument in the previous paragraph. By local incentive compatibility, $T_\tau'(q_\tau(\theta')) = \theta'\nu'(q_\tau(\theta')) < \theta'\nu'(q(\theta')) = T'(q(\theta'))$ for such a consumer since $\nu''(\cdot) < 0$. Then, by continuity, $q_\tau(\theta) > q(\theta)$ and, again since $\nu''(\cdot) < 0$, $T_\tau'(q_\tau(\theta)) = \theta\nu'(q_\tau(\theta)) < \theta\nu'(q(\theta)) = T'(q(\theta))$ for an interval of consumer types containing θ'. Thus, there exists an interval $[\theta_{\min}, \theta_{\max}]$ that contains θ' such that $q_\tau(\theta) \geq q(\theta)$ and $T_\tau'(q_\tau(\theta)) \leq T'(q(\theta))$ for all $\theta \in [\theta_{\min}, \theta_{\max}]$ with strict inequalities for some types.

Since, as argued, the transfer amounts to an expansion in consumers' budget for the seller's good, the menu $\{t(\theta), q(\theta)\}$ is still implementable, so profits must weakly increase. Moreover, the quantity offered to each type in $[\theta_{\min}, \theta_{\max}]$ increases, as just proved, and thus the cost of producing the offered quantities is higher after the transfer. Hence, the price schedule must weakly increase for each type in $[\theta_{\min}, \theta_{\max}]$. □

Proof of Corollary 1: By local incentive compatibility before and after the transfer, $T'(q) = \theta(q)\nu'(q)$ and $T_\tau'(q_\tau) = \theta_\tau(q_\tau)\nu'(q_\tau)$, where $T(q)$ and $T_\tau(q)$ are, respectively, the price schedules before and after the transfer, $q = q(\theta)$, and $q_\tau = q_\tau(\theta)$ with inverse functions $\theta(\cdot)$ and $\theta_\tau(\cdot)$. Differentiating $T_\tau'(q_\tau) = \theta_\tau(q_\tau)\nu'(q_\tau)$ and $T'(q) = \theta(q)\nu'(q)$ gives

$$T_\tau''(q_\tau(\theta)) = \frac{\nu'(q_\tau(\theta))}{q_\tau'(\theta)} + \theta\nu''(q_\tau(\theta)) \text{ and } T''(q(\theta)) = \frac{\nu'(q(\theta))}{q'(\theta)} + \theta\nu''(q(\theta)). \tag{56}$$

Recall from (10) that $G(q) = F(\theta)$ for $q = q(\theta)$ and, similarly, $G_\tau(q_\tau) = F(\theta)$ for $q_\tau = q_\tau(\theta)$. Hence, $f(\theta) = g_\tau(q_\tau(\theta))q_\tau'(\theta) = g(q(\theta))q'(\theta)$, which yields that

$$\frac{1}{q_\tau'(\theta)} = \frac{g_\tau(q_\tau)}{f(\theta)} \leq \frac{g(q)}{f(\theta)} = \frac{1}{q'(\theta)} \Leftrightarrow g_\tau(q_\tau) \leq g(q). \tag{57}$$

Observe that (11) implies that $g_\tau(q_\tau) \leq g(q)$ up to a certain percentile $q_{\max}$ in the distribution of quantities before and after the transfer, where $q_\tau = G_\tau^{-1}(t)$ and $q = G^{-1}(t)$ for $t \in (G_\tau(0), 1)$ by definition of q_τ and q. Hence, $q'(\theta) \leq$

$q'_\tau(\theta)$ up to $q_{\max}$ by (57). The assumption that $q_\tau(\theta) \geq q(\theta)$ for all types further implies that $\nu'(q_\tau(\theta)) \leq \nu'(q(\theta))$ since $\nu''(\cdot) \leq 0$. Given that $\nu'''(\cdot) \leq 0$ by assumption, $\nu''(q_\tau(\theta)) \leq \nu''(q(\theta))$. Thus, it is immediate by (56) that $T''_\tau(q_\tau(\theta)) \leq T''(q(\theta))$ for all percentiles in the distribution of quantities before and after the transfer up to $q_{\max}$. $\square$

Proof of Proposition 5: Rewrite the seller's first-order condition in (9) as

$$\frac{1}{T'(q)} = \frac{1}{c'(Q)} + \frac{[F(\theta) - \gamma(\theta)]}{c'(Q)\theta f(\theta)} = \frac{1}{c'(Q)} + \frac{[G(q) - \gamma(\theta(q))]\theta'(q)}{c'(Q)g(q)\theta(q)} = \frac{1}{c'(Q)} + \frac{[G(q) - \gamma(\theta(q))]\varphi(q)}{c'(Q)g(q)} \tag{58}$$

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

$$\frac{g(q)}{\varphi(q)} \left[\frac{c'(Q)}{T'(q)} - 1\right] = G(q) - \psi(q), \tag{59}$$

where $\psi(q) \equiv \gamma(\theta(q))$. By taking derivatives of each side of (59), 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

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

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. $\square$

Derivation of Reduced Form in (19): The seller's first-order condition in (9) 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 $q'(\theta) = 1/\theta'(q)$. Further manipulating this expression and using the fact that $T'(q) = t_1 p(q)$ by our specification for $T(q)$ yield that

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

which implies

$$\log[p(q)] \approx \log\left[\frac{c'(Q)}{t_1}\right] - \frac{\theta'(q)}{\theta(q)}\left[\frac{1 - \gamma(\theta(q))}{g(q)}\right] + \frac{\theta'(q)}{\theta(q)}\left[\frac{1 - G(q)}{g(q)}\right]. \tag{60}$$

Letting $\gamma(\theta(q)) = \psi(q)$, (60) can be expressed as

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

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

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

and so

$$\begin{aligned}\log[p(q)] \approx\ & \underbrace{f(a,b) - af_1(a,b) - bf_2(a,b) + \frac{a^2}{2}f_{11}(a,b) + \frac{b^2}{2}f_{22}(a,b) + abf_{12}(a,b)}_{\beta_0} \\ &+ \underbrace{[f_1(a,b) - af_{11}(a,b) - bf_{12}(a,b)]}_{\beta_1}\log(q) + \underbrace{[f_2(a,b) - af_{12}(a,b) - bf_{22}(a,b)]}_{\beta_2}\left[\frac{1-G(q)}{g(q)}\right] \\ &+ \underbrace{f_{12}(a,b)}_{\beta_3}\log(q)\left[\frac{1-G(q)}{g(q)}\right] + \underbrace{\frac{1}{2}f_{11}(a,b)}_{\beta_4}\log(q)^2 + \underbrace{\frac{1}{2}f_{22}(a,b)}_{\beta_5}\left[\frac{1-G(q)}{g(q)}\right]^2,\end{aligned}$$

which can equivalently be expressed as

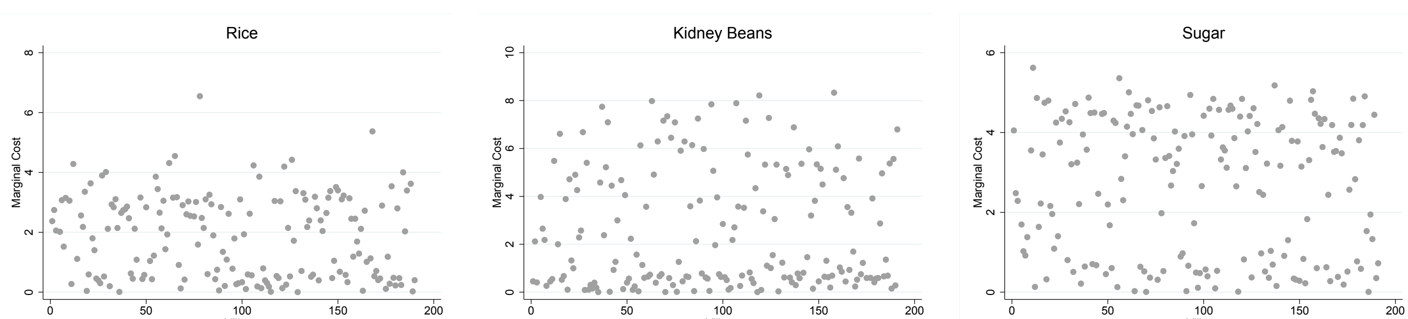
$$\log[p(q)] \approx \beta_0 + \beta_1\log(q) + \beta_2\left[\frac{1-G(q)}{g(q)}\right] + \beta_3\log(q)\left[\frac{1-G(q)}{g(q)}\right] + \beta_4\log(q)^2 + \beta_5\left[\frac{1-G(q)}{g(q)}\right]^2. \quad \square$$

B Appendix: Omitted Estimation Results

We present here estimation results omitted from the main text.[41]

Marginal Cost Estimates. Figure 5 reports the estimated marginal cost at the total quantity provided of each good in each estimated village. The mean of the estimated marginal cost across villages defined as municipalities is 1.724 pesos for rice with a standard deviation of 1.320; 2.396 pesos for kidney beans with a standard deviation of 2.348; and 2.552 pesos for sugar with a standard deviation of 1.709. Hence, although estimated marginal cost varies across villages for each good, its range and mean are similar across goods. Since the villages we study are fairly dispersed and isolated, some variability in estimated marginal cost across villages is to be expected.

Figure 5: Estimated Marginal Cost



Statistics on Estimates. The first three columns of Tables 6 to 8 report the quartiles of the distribution of the t-statistics of the estimates of $c'(Q)$, $\gamma(\theta(q))$, $\log(\theta(q))$, $\log(\nu'(q))$, and $f(\theta)$ across quantities and villages for rice, kidney beans, and sugar. These statistics are meant to illustrate the overall precision of the estimates. The last three

[41] All estimation and counterfactual exercises have been performed in Stata 15 MP.

columns report the quartiles of the distribution across villages of the median t-statistic of these estimates in each village. These statistics are meant to illustrate the variability of the precision of the estimates across villages. As apparent from these tables, the model's primitives as well as the multiplier $\gamma(\theta(q))$ are fairly precisely estimated. Finally, Table 9 reports the quartiles of the distribution of the t-statistics of the estimates of the parameters of equations (15), (16), and (17) across villages. These equations correspond, respectively, to the estimated specification of $T(q)$, $G(q)$, and the seller's first-order condition in each village. As apparent from the table, these parameters are also fairly precisely estimated.

Table 6: Distribution of t-statistics of Estimates for Rice

	Overall			Between-Village Quartiles of Village Median		
	p_{25}	p_{50}	p_{75}	p_{25}	p_{50}	p_{75}
$c'(Q)$	4.649	18.902	40.711	4.649	18.902	40.711
$\gamma(\theta(q))$	49.845	280.000	2321.956	83.446	314.776	1066.796
$\log(\theta(q))$	1.687	4.241	12.226	2.385	4.706	12.350
$\log(\nu'(q))$	1.085	3.026	9.073	1.707	3.637	7.817
$f(\theta)$	4.114	10.579	16.008	7.500	10.368	14.405

Table 7: Distribution of t-statistics of Estimates for Kidney Beans

	Overall			Between-Village Quartiles of Village Median		
	p_{25}	p_{50}	p_{75}	p_{25}	p_{50}	p_{75}
$c'(Q)$	3.512	8.188	26.696	3.512	8.188	26.696
$\gamma(\theta(q))$	71.301	255.191	945.454	112.660	255.065	612.843
$\log(\theta(q))$	1.158	3.835	8.643	1.867	4.550	8.781
$\log(\nu'(q))$	0.664	2.597	10.226	1.084	3.420	8.464
$f(\theta)$	4.330	10.548	17.428	8.790	11.299	17.035

Table 8: Distribution of t-statistics of Estimates for Sugar

	Overall			Between-Village Quartiles of Village Median		
	p_{25}	p_{50}	p_{75}	p_{25}	p_{50}	p_{75}
$c'(Q)$	9.315	37.044	102.045	9.315	37.044	102.045
$\gamma(\theta(q))$	90.271	267.558	754.383	114.995	249.198	532.734
$\log(\theta(q))$	2.416	5.824	12.896	3.286	7.118	13.931
$\log(\nu'(q))$	1.427	5.323	22.199	2.353	8.163	20.584
$f(\theta)$	5.667	11.619	18.337	8.840	12.361	17.607

Table 9: Distribution of t-statistics of Parameters by Good

	Rice			Kidney Beans			Sugar		
	p_{25}	p_{50}	p_{75}	p_{25}	p_{50}	p_{75}	p_{25}	p_{50}	p_{75}
t_0	230.776	387.465	633.870	419.496	668.494	1203.940	488.846	812.356	1459.720
t_1	1.798	4.017	7.829	10.711	25.837	46.734	24.227	40.628	68.174
g_0	14.884	30.924	52.712	2.199	3.755	6.112	3.662	7.533	12.846
g_1	13.273	30.277	52.464	0.397	1.503	3.502	1.221	3.031	20.417
g_2	4.759	15.067	34.319	1.221	1.709	2.677	1.706	2.839	3.802
g_3	0.943	4.525	19.341	13.920	13.920	13.920	35.639	35.639	35.639
g_4	12.631	26.762	45.275	1.991	3.116	4.545	7.015	9.717	15.506
g_5	8.477	17.251	35.139	6.000	13.116	20.780	29.268	40.037	62.714
c	2.280	11.186	74.364	1.512	4.442	19.274	5.414	19.831	75.613
d_0	5.245	11.914	37.571	2.290	4.422	14.124	3.270	8.169	30.075
d_1	2.784	8.973	37.825	4.606	8.923	29.260	6.206	14.297	38.293
a_0	13.815	32.662	106.822	6.216	15.209	41.150	6.615	19.618	70.537
a_1	7.431	19.279	1187.317	2.099	8.861	52.094	8.461	15.658	43.505
a_2	17.552	49.238	101.920	4.453	11.238	30.227	5.446	12.633	26.724