Nonlinear Pricing in Village Economies*

Orazio Attanasio
Yale University, IFS, FAIR at NHH, NBER, and CEPR

Elena Pastorino
Hoover Institution, Stanford University,
and Federal Reserve Bank of Minneapolis

Economics Working Paper 20109

HOOVER INSTITUTION
434 GALVEZ MALL
STANFORD UNIVERSITY
STANFORD, CA 94305-6010

January 2020

This article was originally published in *Econometrica*. The citation for the article is: "Nonlinear Pricing in Village Economics," by Orazio Attanasio and Elena Pastorino, *Econometrica*, Vol. 88, No.1 (Jan.2020). Copyright © 2020 by The Econometric Society, 2020. https://doi.org/10.3982/ECTA13918

* Orazio Attanasio: orazio.attanasio@yale.edu
Elena Pastorino: epastori@stanford.edu
We are grateful to Aureo de Paola, Bruno Jullien, 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 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.

Nonlinear Pricing in Village Economies
Orazio Attanasio and Elena Pastorino
Economics Working Paper 20109
January 2020
Keywords: Price discrimination, heterogeneous budget constraints, heterogeneous reservation utilities, countervailing incentives, calorie constraints, subsistence constraints, cash transfers, Progresa, Mexico, identification, structural estimation.

Orazio Attanasio
Yale University,
IFS, FAIR at NHH,
NBER, and CEPR

Elena Pastorino
Hoover Institution,
Stanford University, and Federal
Reserve Bank of Minneapolis

Abstract:
This paper examines the prices of basic staples in rural Mexico. We document that nonlinear pricing in the form of quantity discounts is common, that quantity discounts are sizable for basic staples, 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 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 on three commonly consumed commodities from municipalities and localities in Mexico. Interestingly, we find that relative to linear pricing, nonlinear pricing is beneficial to a large number of households, including those consuming small quantities, mostly because of the higher degree of market participation that nonlinear pricing induces. We also show that the Progresa transfer has affected the slopes of the price schedules of the three commodities we study, which have become steeper as consistent with our model, leading to an increase in the intensity of price discrimination. Finally, we find 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 in a village, accounts for the shift in price schedules induced by the program.

1. INTRODUCTION

QUANTITY DISCOUNTS in the form of unit prices declining with quantity are pervasive in developing countries. McIntosh (2003), for instance, documented differences in the price of drinking water paid by poor and rich households in the Philippines. Pannarunothai and Mills (1997) and Fabricant, Kamara, and Mills (1999) reported similar differences in the price of health care and services in Thailand and Sierra Leone. Attanasio and Frayne (2006) showed evidence that households purchasing basic staples in Colombian villages face price schedules rather than linear prices: richer households buy larger quantities of the same goods that poorer households purchase, but richer households 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 a seller's ability 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. Hence, consumers of smaller quantities, who are typically the poorest ones in developing countries, tend to suffer greater distortions relative to consumers of larger 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. The standard model therefore 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. By affecting consumption, any such realistic dimension of heterogeneity across consumers may naturally have important consequences for the welfare implications of any pricing scheme.

Sellers' 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 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 prices of commodities but have consistently found no impact. For instance, Hoddinott, Skoufias, and Washburn (2000) concluded that "there is no evidence that Progresa communities paid higher food prices than similar control communities" (p. 33). Similarly, when Angelucci and De Giorgi (2009) assessed the impact of Progresa on the consumption of noneligible households, they dismissed the possibility that their results are mediated by changes in local unit prices. Although Progresa has not affected average unit prices, in the presence of nonlinear pricing, the program may nonetheless have resulted in differential changes in the unit prices of different quantities and may have therefore had 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

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

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, which vary across consumers. We characterize nonlinear pricing in this model and investigate the effect 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 sizable quantity discounts for common staples. We also find that Progresa has had a significant effect on unit prices by leading to an increase in the magnitude of quantity discounts, namely, to steeper schedules of unit prices, but that this effect cannot be detected without accounting for the dependence of unit prices on the quantities purchased.

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 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 establishing a formal equivalence between these 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 marginal prices and marginal costs depends on the difference between the cumulative multiplier associated with consumers' participation or budget constraints and the cumulative distribution function of consumers' characteristics. This relationship 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 with data on three commodities, namely, rice, kidney beans, and sugar, from 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 hazard rate and reverse hazard rate of the distribution of consumers' 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: the unit prices of smaller quantities are higher than the unit prices of larger quantities. Like previous studies, we estimate that the Progresa transfer has not affected average unit prices. However, we find that the transfer has increased the (absolute value of the) slope of unit prices with respect to quantity and so has led to an increase in the intensity of price discrimination, which has affected both households beneficiaries of the program and noneligible households. Finally, using an approximate reduced form of the optimality conditions for seller and consumer behavior, we show that our model can account for the change in price schedules induced by Progresa.

The intuition behind some of our theoretical results, in particular those about the effect of nonlinear pricing on market participation and about the impact of income transfers on prices, is simple and deserves a mention. In terms of market 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 with consumers who have access to especially attractive outside options. Such consumers would be excluded from the market under linear pricing. The argument is as follows. To induce such 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 substantially lower profits from all existing consumers. Thus, including such 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 like the Progresa one that are more generous for poorer households, which tend to purchase smaller quantities, can intensify price discrimination, thereby exacerbating some of the distortions associated with nonlinear pricing.

Our empirical results are consistent with these intuitions. Our estimates imply that sellers have market power in the villages in our data and exercise it by price discriminating across consumers through distortionary quantity discounts. Interestingly, a substantial fraction of consumers, including those of small quantities, consume *above* the first best rather than, as often argued and implied by the standard model, below it. We then compare observed nonlinear pricing to a counterfactual scenario in which sellers have market power but 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 the unit prices of rice, kidney beans, and sugar, we explicitly account for their variation across quantities and allow the program to affect their entire schedules. 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 schedules of unit prices of the three goods have become significantly *steeper* after transfers have been introduced, with greater discounts for large quantities in villages receiving the Progresa transfer. Namely, the transfer has led to an increase in the intensity of price discrimination. We also derive a reduced form of the model from consumer and seller optimality conditions that relates unit prices to both 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 price schedules induced by the program, in particular the change in their slopes, arises from a shift in the hazard rates of the distributions 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 impact 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 data set 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, De Giorgi, and Jayachandran (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 focused on 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 the ages of 9 and 16 and was conditioned on regular school attendance. Although the program administration was relatively strict in enforcing these conditionalities, they were not likely to be 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. Specifically, 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 506 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 in these localities in December 1999. Households in these localities were followed for several periods. In our empirical exercise, we use the surveys of October 1998, March and November 1999, November 2000, and 2003. We could not use some waves, such as those collected in 1997 and March 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; food accounts for a substantial share of their consumption. However, not all households within 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 exhibits substantial variation not only within but also across villages. This heterogeneity is reflected in the variability of the rate of eligible households across villages, for instance.

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. We will exploit this variation when examining key implications of the model we propose, in particular about the impact of cash transfers on prices. Second, the data provide a census of 506 villages in that all households in the relevant localities are surveyed, thereby allowing us to estimate the entire distribution of quantities purchased and prices paid in each village, at least for commodities that are commonly bought. Third, the data are very rich, as we now detail.

The consumption and expenditure module of the surveys contains information crucial for the purpose of our paper. Each household is interviewed and asked to report not only the quantity consumed of 36 food commodities during the week preceding an 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, and meat and other animal products, and are supposed to be exhaustive of the foods consumed by households. In what follows, we focus on commodities that are relatively homogeneous in their quality and are purchased and consumed by most households, as explained below.

Given the information available on expenditures and quantities purchased for each recorded item, it is possible to compute their *unit values*, as defined by the ratios of these variables. From now on, we refer to unit values as *unit prices*. Attanasio, Di Maro, Lechene, and Phillips (2013) discussed 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 found that average and median unit values well approximate unit prices collected from local stores, which are available for some commodities in the locality surveys. They also found that unit values closely match national data on prices.

[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.

Naturally, differences in the rate of households eligible for Progresa across villages are related to differences in the distribution of quantities purchased across villages. Such heterogeneity could account also for differences in unit prices across villages. Furthermore, the variability in the rate of eligible households across villages is likely to affect how the Progresa transfer has modified the distribution of consumption and, according to the model we propose, unit prices in different villages. As we will show in Section 5.4, changes in the distributions of quantities purchased *within* villages in response to the Progresa transfer are key to assessing the ability of our model to account for differences in price schedules *across* villages and so for the impact of the program on unit prices. We now turn to describe the impact of the Progresa transfer on unit prices.

Quantity Discounts and Price Effects of Progresa. Quantity discounts are common in several markets in developing countries. Attanasio and Frayne (2006), for instance, estimated the supply schedule for several basic food staples, including rice, carrots, and beans, in Colombian villages and documented substantial discounts for large volumes. Specifically, they found that the elasticity of the unit price of rice with respect to the quantity purchased is as large as 0.11 in absolute value in their preferred specification. They estimated even larger quantity discounts for different specifications and other goods such as carrots or beans.

Here, we first document the existence of patterns of quantity discounts 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 our 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 not observing any purchases for a household does not necessarily reflect exclusion from the market but could be due simply to the timing of the Progresa interviews. Hence, the assumption of full market participation we will formulate in our analysis is not implausible. Indeed, across localities in the week preceding an interview, the median fraction of households consuming rice 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 in-kind payment: the median fraction of households that purchase the amount consumed of each of these goods across localities is 100% for rice, 94% for kidney beans, and 100% for sugar.

We use data from the Progresa waves of October 1998, March 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 we consider, we examine the relationship between unit prices and quantities purchased in each village.

Columns 1 in Table I contain estimates from a regression of log (real) unit prices on a constant and log quantities. The different numbers of observations in each row reflect

[3]Some studies have argued that the dispersion in the unit price of a good observed in a market might reflect differences in quality. Deaton (1989), for instance, argued 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 unit price schedules, contrary to what we observe.

TABLE I

PRICE SCHEDULES AND IMPACT OF CASH TRANSFERS ON PRICES (98% TRIMMING)[a]

	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

[a]Note: Wave fixed effects are included. Standard errors are clustered at the locality level.

the different numbers of purchases we observe in our sample. In this exercise as well as in those in columns 2 and 3, we include wave fixed effects and cluster standard errors at the locality level. We estimate that the elasticity of unit prices with respect to quantity 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 unit prices on a constant and a dummy, "Treatment," equal to 1 for transactions occurring in localities targeted by the program. Consistent with studies that have estimated this impact, such as Hoddinott, Skoufias, and Washburn (2000), we do not find any evidence that the Progresa transfer has affected the average unit prices of the three goods.

We complement this evidence on the impact of the program by examining the possibility that the Progresa transfer has 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 of each good has increased in absolute value with the program: it has changed 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% level 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, which we will further examine in Section 5.4. See the Supplemental Material (Attanasio and Pastorino (2020)) for analogous results when quantities and real expenditures are trimmed at the top 1% or 5%, rather than at the top 2%.

Market Structure. The model we will 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, one would ideally like to consider a relatively isolated market with one seller or a small number of them.[4] 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 localities and the number of recorded purchases. Moreover, despite some localities being quite isolated, from an administrative point of view all localities belong to a municipality and are often connected in several ways. For instance, it is not unusual for some households to shop for certain items in a locality within the municipality of residence but different from the locality where they live. For these reasons, in the main text we focus on villages defined as municipalities. However, we estimate our model on villages defined as both municipalities and localities and obtain fairly similar results for these two definitions of villages, as discussed in Section 5.

Note that 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 data set, 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 municipalities, 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.

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; or they may receive a good as a transfer from relatives, friends, or the government. As the desirability or feasibility of these alternative consumption possibilities may differ across consumers, so does consumers' reservation utility.

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

[4] Although we focus on the problem of a single seller, the model we develop can account for different degrees of seller market power by interpreting a consumer's reservation utility as the utility obtained when purchasing from other sellers. 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 Supplemental Material.

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 budget constraints. As consumers typically have preferences for multiple goods, we allow for consumers' substitution across them and model subsistence constraints as arising from calorie constraints that 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 $f(\theta)$ with $f(\theta) > 0$ 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, which can nevertheless entail 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 will relax this assumption in the empirical analysis. We denote by $s(\theta, q) = v(\theta, q) - c(q)$ the *social surplus* from quantity q. We assume that $s_q(\theta, q)/v_{\theta q}(\theta, q)$ decreases with q, which ensures that the seller's problem admits a

[5]The optimal price schedule is known for special cases, for instance, when 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.

[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 Tadelis and Segal (2005).

unique solution and that first-order conditions are necessary and sufficient to characterize it. This assumption plays the same role as the assumptions in the standard model that $s(\cdot,\cdot)$ is concave in q and $v_\theta(\cdot,\cdot)$ is convex in q. We define the *first-best* quantity, $q^{\mathrm{FB}}(\theta)$, 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 1—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 consumers not to participate, then the equilibrium price schedule faced by consumer 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 (IC) and participation or individual rationality (IR) constraints, that is,

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

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

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

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 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 interchangeably use these *menu* and *tariff* interpretations throughout.

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

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

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

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

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) formulated three important assumptions to characterize an optimal menu for the IR problem: *potential separation* (PS), *homogeneity* (H), and *full participation* (FP). Consider (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 $\gamma(\theta)$, namely, $q(\theta) = l(\gamma(\theta), \theta)$. 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 $\{\overline{u}(\theta), \overline{q}(\theta)\}$ with full participation is implementable. This assumption ensures that a consumer's IC constraint can be satisfied when the IR constraint binds.[10] Assumption (FP) states that all types participate. Sufficient conditions for (FP) are (H) and $s(\theta, \overline{q}(\theta)) \geq \overline{u}(\theta)$ for all types. This latter condition guarantees that the seller has an incentive to trade with all consumers.

Jullien (2000) showed 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 $q(\theta) = l(\gamma(\theta), \theta)$ for each consumer type with associated price $t(\theta)$ and utility $u(\theta)$, which equals $\overline{u}(\theta)$ for consumer types whose IR constraints bind. Note that since $q(\theta)$ is continuous, $\gamma(\theta)$ can have mass points only at $\underline{\theta}$ or $\overline{\theta}$, and thus the IR constraints can bind at isolated points only for $\underline{\theta}$ or $\overline{\theta}$. When $\gamma(\theta) = 1$ for all types so that the IR constraints bind only for $\underline{\theta}$, 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 lowest degree under perfect competition to the highest one under monopoly. Specifically, when the reservation utility equals the social surplus under the first best for each type, that is, $\overline{u}(\theta) = v(\theta, q^{\mathrm{FB}}(\theta)) - c(q^{\mathrm{FB}}(\theta))$, the solution to the seller's problem implies $\gamma(\theta) = F(\theta)$ for all consumers so that consumers purchase first-best quantities at cost from the seller. As the reservation utility is lowered from this maximal value for each type, profits correspondingly increase, thus allowing the model to capture any degree of imperfect competition. This feature of the model provides an important dimension of flexibility relative to the standard model for the measurement exercises in later sections.

[8] See the Supplemental Material 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.

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

[10] With $v_\theta(\theta, q) > 0$ by assumption, (H) implies that $\overline{u}(\theta)$ is monotone since it requires $\overline{u}'(\theta) = v_\theta(\theta, \overline{q}(\theta))$. (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 competitor; see the Supplemental Material for details. 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 $-\theta$.

3.2. *A Model With Heterogeneous Budget Constraints*

Suppose now that instead of having access to heterogeneous outside options, consumers face heterogeneous subsistence constraints. We show that these constraints limit the amount of resources that a consumer can spend on a seller's good and formally give rise to a *budget constraint* for the good. We also establish 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. In the next subsection, we will use this model to examine the impact of income transfers on prices and consumption.

Setup. Assume 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 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, respectively, 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 from the seller, the consumer's problem is

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

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

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

where $I(\theta, q, w)$ is the maximal amount that the consumer can 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 with θ and q. An intuition for why $I(\theta, q, w)$ may increase with θ, and thus $\underline{z}(\theta, q)$ may decrease with θ, is that a consumer may have a greater taste for the

[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 the constraint $C^q q + C^z z \geq \underline{C}$, which is considered, for instance, by Jensen and Miller (2008), where C^q and C^z are the calories provided by one unit of the seller's good and one unit of the numeraire, respectively, and $\underline{C}$ is the subsistence intake.

[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.

seller's good precisely because the consumer derives more calories from consuming it, and thus requires less of other goods to achieve a given calorie intake. The requirement that $I(\theta, q, w)$ weakly increase with q, and so $\underline{z}(\theta, q)$ weakly decrease with q, is equivalent to $C^q(\theta, q)$ weakly increasing with q and is natural: the greater the amount of the seller's good consumed, the greater the calorie intake, and so the smaller the amount of z necessary to meet the calorie constraint. 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 maximizes expected profits subject to consumers' incentive compatibility (IC), participation or individual rationality (IR'), and budget (BC) constraints, that is,

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

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

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

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

We refer to this model in which the seller's constraints are IC, IR', and BC as the *BC model* and define an allocation $\{u(\theta), q(\theta)\}$ that satisfies them as *implementable*. Although we allow for heterogeneity among consumers in both θ and w, in this section we focus on 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 will 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 with assumption (H), we assume that there exists an incentive compatible menu $\{\overline{t}(\theta), \overline{q}(\theta)\}$ that induces each consumer to purchase and spend her entire budget for the seller's good, that is,

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

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 $\{\overline{t}(\theta), \overline{q}(\theta)\}$ that induces all consumers to participate.[14]

Since income affects consumers' purchase behavior, changes in the distribution of consumers' income arising, for instance, from income transfers typically influence a seller's

[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 for type θ. Assumption (BCH) and $I(\theta, \overline{q}(\theta)) \geq c(\overline{q}(\theta))$ for all types 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 such a consumer. Thus, all types participate.

optimal 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 income. Our equivalence result between the BC model and the IR model implies that the BC model can be viewed as providing a map between changes in income and changes in reservation utility. In this sense, the IR model can be considered as a "reduced form" of the BC model. We will explore the implications of the BC model for the impact of income transfers in the next subsection, when we analyze the effect of Progresa on prices and consumption.

Equivalence Between Participation and Budget Constraints. The seller's problem with IC, IR', and BC constraints has no known solution. Here, we proceed to characterize the seller's optimal menu indirectly by establishing an equivalence between the IR problem and the BC 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 $I(\theta, \hat{q}(\theta)) = v(\theta, \hat{q}(\theta)) - \overline{u}_{\text{IR}}(\theta)$ for any allocation $\{\hat{u}(\theta), \hat{q}(\theta)\}$ in the BC model, where $\overline{u}_{\text{IR}}(\theta)$ denotes the reservation utility of a consumer of type θ in the IR model. Since $\hat{t}(\theta) = v(\theta, \hat{q}(\theta)) - \hat{u}(\theta)$, it is immediate that the BC constraint of the BC problem is equivalent to the IR constraint of the IR problem in this case. Although this approach is intuitive since it directly relates reservation utilities to budgets, it is unduly restrictive: it requires the schedules of reservation utilities in the IR problem and budgets in the BC problem to agree for each type at any 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 budget schedule with respect to quantity, agree just for types whose IR constraints bind at the optimal allocation for the IR problem—as long as consumers have enough income in the BC problem to afford the IR allocation.

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

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

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

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

for each type, along with the complementary slackness condition

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

By Result 1 in the proof of Proposition 1, 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 $\Phi(\overline{\theta}) = 1$.[15] Denote by $\{t_{\mathrm{IR}}(\theta), q_{\mathrm{IR}}(\theta)\}$ the optimal menu, by $\{u_{\mathrm{IR}}(\theta), q_{\mathrm{IR}}(\theta)\}$ the optimal allocation, and by $\{\overline{u}_{\mathrm{IR}}(\theta), \overline{q}_{\mathrm{IR}}(\theta)\}$ the reservation utility and quantity profiles in the IR model. We now establish the desired equivalence.

PROPOSITION 1: *Suppose that the allocation that solves the IR problem is affordable in the BC problem in that $I(\theta, q_{\mathrm{IR}}(\theta)) \geq v(\theta, q_{\mathrm{IR}}(\theta)) - \overline{u}_{\mathrm{IR}}(\theta)$, with equality for consumer types whose IR constraints bind, and $\overline{u}_{\mathrm{IR}}(\underline{\theta}) \geq \overline{u}$. If $I_q(\theta, q_{\mathrm{IR}}(\theta))$ equals $v_q(\theta, q_{\mathrm{IR}}(\theta))$ for types whose IR constraints bind, then the solution to the IR problem solves the BC 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 that a seller can charge to a consumer, since the requirement $u_{\mathrm{IR}}(\theta) \geq \overline{u}_{\mathrm{IR}}(\theta)$ is equivalent to $t_{\mathrm{IR}}(\theta) \leq v(\theta, q_{\mathrm{IR}}(\theta)) - \overline{u}_{\mathrm{IR}}(\theta)$, 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 $v_q(\theta, q_{\mathrm{IR}}(\theta))$ equals $I_q(\theta, q_{\mathrm{IR}}(\theta))$ for consumers whose IR constraints bind in the IR problem and so whose BC constraints bind ($\phi(\theta) > 0$) in the BC problem. 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 that $I_q(\theta, q_{\mathrm{IR}}(\theta)) = v_q(\theta, q_{\mathrm{IR}}(\theta))$ when the IR constraints bind. This condition implies simply that if the seller uses the budget schedule $I(\theta, \overline{q}_{\mathrm{IR}}(\theta))$ as a price schedule in the BC model when the BC constraints bind, then he can induce consumers to demand the same incentive compatible quantities that they demand in the IR model when the IR constraints bind. Hence, consumers with binding constraints in the two models can be induced to purchase the same quantities.[16]

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 that directly affect consumers' budgets, such as income transfers. We do so in the next subsection.

[15]Like Jullien (2000), we presume that these conditions are necessary and sufficient to characterize a solution to the BC problem. An alternative equivalence argument can be derived, for instance, by applying Theorem 1 in Jullien (2000) to a version of the BC problem in which the BC constraints are replaced by the constraints $v(\theta, q(\theta)) - t(\theta) \geq \overline{u}_{BC}(\theta)$ for any θ, where $\overline{u}_{BC}(\theta) \equiv v(\theta, \overline{q}(\theta)) - I(\theta, \overline{q}(\theta))$ and $\overline{q}(\theta)$ is defined in assumption (BCH).

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

3.3. *Properties and Implications of Nonlinear Pricing*

In light of 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. Here, we first 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 face 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 the optimal quantity schedule and the reservation quantity schedule are increasing with the type.[17] Hereafter, results apply to consumer types in $(\underline{\theta}, \overline{\theta})$ when $f(\theta)$ is not strictly positive everywhere.

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

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

The price schedule $T(q)$ exhibits *quantity discounts* if $T''(q) \leq 0$ or the unit price $p(q) = T(q)/q$ declines with q.[18] Denote by $A(q) \equiv -\nu''(q)/\nu'(q)$ the coefficient of absolute risk aversion evaluated at q. Observe that $[1 - F(\theta)]/f(\theta)$ is the inverse of the hazard rate and $F(\theta)/f(\theta)$ is the inverse of the reverse hazard rate of the distribution of consumer types. We can prove the following result.

PROPOSITION 2: *Assume that $\nu''(\cdot) < 0$ and $d([1 - F(\theta)]/f(\theta))/d\theta \leq 0$. If $F(\theta)/[\theta f(\theta)] \leq \min\{1, d(F(\theta)/f(\theta))/d\theta\}$ and $\overline{u}''(\theta) \geq \nu'(\overline{q}(\theta))/[\theta A(\overline{q}(\theta))]$ for each $\theta \in (\underline{\theta}, \overline{\theta})$, then $T''(q) \leq 0$ for each $q = q(\theta)$ and $\theta \in (\underline{\theta}, \overline{\theta})$.*

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 $F(\theta)/f(\theta)$. The restriction that $F(\theta)/f(\theta) \leq \theta$ simply bounds the rate of increase of $T(q)$ when a seller offers quantities above the first best, as is the case when $\gamma(\theta) < F(\theta)$. Intuitively, quantity discounts in the form of $T''(q) \leq 0$ require the rate of increase of $T(q)$ to decrease with quantity. The additional restriction that $d(F(\theta)/f(\theta))/d\theta \geq F(\theta)/[\theta f(\theta)]$ strengthens the usual condition on the distribution

[17] For instance, the optimal quantity schedule increases with θ if, as assumed, $s_q(\theta, q)/v_{\theta q}(\theta, q)$ decreases with q and, as consistent with (PS), $s_q(\theta, q)/v_{\theta q}(\theta, q)$ increases with θ, $F(\theta)/f(\theta)$ increases with θ, and $[1 - F(\theta)]/f(\theta)$ decreases with θ. Of the last three monotonicity conditions, it is sufficient that either the first or the last two be strict. Note that $v_\theta(\theta, q) > 0$ implies that $\nu(q) > 0$.

[18] A sufficient condition for $p'(q) \leq 0$ is indeed $T''(q) \leq 0$ 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 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 $f(x_2)/x_2$ decreases with quantity.

of consumer types for assumption (PS) to hold in models with heterogeneous reservation utilities, namely, $d(F(\theta)/f(\theta))/d\theta \geq 0$. It guarantees that a seller has an incentive to discriminate across consumers.[19] Consider now the condition on $\overline{u}''(\theta)$, which requires it to be large enough. This condition ensures that consumers whose IR constraints bind are offered quantity discounts. In general, the convexity of $\overline{u}(\cdot)$ implies that outside consumption opportunities are increasingly more valuable for consumers of higher types. Since $\overline{u}'(\theta) = v(\overline{q}(\theta))$ by assumption (H), by offering larger quantities at lower marginal prices, a seller can satisfy higher types' IR constraints and induce them to buy more than lower types, thereby separating higher types from lower ones. Quantity discounts are then optimal for a seller.

Consider now the model's implications for consumption. By comparing the first-order condition in (9) with that for the first-best quantity, $T'(q(\theta)) = c$, it is immediate that the quantity provided to a consumer of type θ is below the first best when $\gamma(\theta) > F(\theta)$ and 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 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 arises instead 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. A seller can do so while separating higher types from lower ones by offering quantities meant for higher types that are *above* higher types' first-best level of consumption at marginal prices below marginal cost. By doing so, a seller can not only induce higher types to purchase these large quantities but also distinguish them from lower types, who naturally prefer smaller quantities. Then, a seller can differentiate consumers because lower types would need to purchase much larger quantities than is desirable to them to imitate the behavior of higher types.[20]

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. Conditional on purchasing the good, a consumer of type θ chooses the quantity $q_m(\theta)$ and obtains utility $u_m(\theta) = \theta v(q_m(\theta)) - p_m q_m(\theta)$. (Formally, $u_m(\theta)$ is the value of the consumer's problem under linear pricing once the participation constraint is dropped, and $q_m(\theta)$ satisfies $\theta v'(q_m(\theta)) = p_m$.) It turns out that when all consumers participate under both pricing schemes and nonlinear pricing entails quantity discounts, consumers are better off under linear pricing in that $u_m(\theta) \geq u(\theta)$. 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,

[19] All these conditions on the distribution of types are satisfied, for instance, by a uniform distribution and a four-parameter beta distribution with shape parameters $\alpha \geq 1$ and $\beta = 1$.

[20] Since $\gamma(\theta) \leq 1$, the augmented model gives rise to (weakly) higher levels of consumption and, correspondingly, (weakly) 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 is (weakly) higher in the augmented model that in the standard model for each type, consumer surplus is clearly (weakly) higher in the augmented model.

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 outside consumption opportunities in that $\overline{q}(\theta) > q^{\mathrm{FB}}(\theta)$ can be excluded under linear pricing and so are better off under nonlinear pricing.

PROPOSITION 3: *The following results hold*:

(1) *Assume that* (*FP*) *holds under linear pricing. If either* (*a*) $p'(q) \leq 0$ *at* $q = q(\theta)$ *and* $q_m(\theta) \geq q(\theta)$ *or* (*b*) $T''(q) \leq 0$ *at* $q \geq q(\theta)$, $q(\theta) \geq q_m(\theta)$, *and* $\gamma(\theta) < 1$, *then a consumer of type* θ *is better off under linear pricing*.

(2) *Let* $\nu''(\cdot) < 0$. *Assume that* $\overline{q}(\theta) > q^{\mathrm{FB}}(\theta)$ *for consumer types in* $[\theta', \theta'']$. *If there exists a type* $\hat{\theta}$ *in* $[\theta', \theta'']$ *with* $u_m(\hat{\theta}) = \overline{u}(\hat{\theta})$, *then some consumer types in* $(\hat{\theta}, \theta'']$ *are excluded under linear pricing and so are better off under nonlinear pricing*.

To understand the role of the condition $\overline{q}(\theta) > q^{\mathrm{FB}}(\theta)$ in part (2) of Proposition 3, note that since $\overline{u}'(\theta) = \nu(\overline{q}(\theta))$ by assumption (H), large values of $\overline{q}(\theta)$ are associated with a rapidly increasing reservation utility profile.[21] To induce consumers with especially attractive outside consumption possibilities to participate, a seller must 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 greatly lower profits from all existing consumers. 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 marginal prices, a seller may have an incentive to serve those consumers who would be unprofitable under linear pricing. In Section 5.3, we will examine the extent to which this implication of our model is borne out in the data. See the Supplemental Material for an illustration of Proposition 3.

Income Transfers. Here, we show that when consumers face a budget constraint for a seller's good, income transfers increase consumption but also lead typically to an increase in prices, as a seller adjusts offered quantities and prices in response to consumers' greater ability to pay.[22] Intuitively, when consumers are constrained by a budget for a seller's good, an increase in their income affects prices by creating an incentive for a 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 consumers' behavior. Indeed, a seller maximizes profits by increasing the price $T(q)$ of each quantity q by the amount of the transfer.

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

[22] In light of assumption (FP), it is implicit in this comparative static exercise that the (IR') constraints are satisfied before and after the transfer.

Now consider the 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). Since poorer households consume less of the normal goods we consider than richer households and our model implies a monotone relationship between types and quantities consumed, assuming that such a transfer satisfies $\tau'(\theta) \leq 0$ then seems consistent with the data.

To understand the impact of a progressive transfer $\tau(\theta)$ with $\tau'(\theta) \leq 0$, recall that a consumer of type θ pays $t(\theta)$ to purchase $q(\theta)$ from the seller, and spends 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 $(t_\tau(\theta), q_\tau(\theta))$. Therefore, as in the case of a uniform transfer, the seller can ask for a higher price without excluding any consumer. Unlike in the case of a uniform transfer, though, since consumers' ability to pay increases differentially with the transfer, the seller has an incentive to charge different consumer types different marginal prices after the transfer. In particular, it turns out that if $\tau'(\theta) < 0$ and the budget for the seller's good satisfies $I_{\theta q}(\cdot, \cdot) \geq 0$, then any consumer who spends her entire budget on the seller's good before and after the transfer demands a larger quantity. To preserve incentive compatibility, a seller must then offer larger quantities, and correspondingly 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 at least some consumers increases by an argument analogous to that in the case of a uniform transfer.

PROPOSITION 4: *Let $v''(\cdot) < 0$. Assume that $\tau(\theta) \geq 0$ and $\tau'(\theta) \leq 0$ for all consumer types with both sets of inequalities strict for at least an interval Θ' of types. Suppose that the budget constraint binds before and after the transfer for at least type θ' in Θ', that the optimal menus before and after the transfer satisfy the conditions of Proposition 1, and that $I_{\theta q}(\cdot, \cdot) \geq 0$. Then, the transfer leads to greater consumption and a higher price schedule with lower marginal prices for all types in a subinterval of Θ' that includes θ'.*

When both $T(q)$ and q increase, 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 argue, 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 $\tau(\theta)$ on the nonlinearity of prices, we compare the curvature of the price schedule before and after the transfer is introduced type by type or, equivalently, at the same percentiles in the distributions of quantities purchased before and after the transfer. Since optimal quantity profiles are monotone, a given percentile in the two quantity distributions corresponds to the same type.

To see this point formally, denote by $\{q_\tau(\theta)\}$ the quantity profile after the transfer is introduced and by $G(q)$ and $G_\tau(q_\tau)$, respectively, the cumulative distribution functions of quantities purchased before and after the transfer with associated probability density functions $g(q)$ and $g_\tau(q_\tau)$, where $q = q(\theta)$ and $q_\tau = q_\tau(\theta)$. For any quantity $q = q(\theta)$ purchased before the transfer by a consumer of type θ, the corresponding quantity purchased after the transfer by the same consumer is $q_\tau = q_\tau(\theta)$. Then, the associated πth percentile in the two distributions of quantities purchased before and after the transfer satisfies $\hat{\pi} = G_\tau(q_\tau) = G(q)$ with $\hat{\pi} = \pi/100$ for $q_\tau = q_\tau(\theta)$ and $q = q(\theta)$, since

$$\begin{aligned}\hat{\pi} &= G_\tau(q_\tau) = \Pr\bigl(q_\tau(\tilde{\theta}) \le q_\tau\bigr) = \Pr\bigl(\tilde{\theta} \le q_\tau^{-1}(q_\tau) = \theta\bigr) \\ &= F(\theta) = \Pr\bigl(\tilde{\theta} \le q^{-1}(q) = \theta\bigr) = \Pr\bigl(q(\tilde{\theta}) \le q\bigr) = G(q). \end{aligned} \tag{10}$$

Consider now a transfer that increases consumption in that

$$g_\tau\bigl[G_\tau^{-1}(\hat{\pi})\bigr] \le g\bigl[G^{-1}(\hat{\pi})\bigr] \text{ for all } \hat{\pi} \in \bigl(G_\tau(0), \hat{\pi}_{\max}\bigr). \tag{11}$$

Condition (11) states that 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 in the two distributions of quantities up to the $\pi_{\max}$th percentile. This percentile in the distribution of quantities purchased before the transfer is quantity $q(\theta_{\max})$ and in the distribution of quantities purchased after the transfer is quantity $q_\tau(\theta_{\max})$ for some type $\theta_{\max}$. By (11), up to the $\pi_{\max}$th percentile, the distribution function $G_\tau(\cdot)$ assigns less mass to smaller quantities than the distribution function $G(\cdot)$. Indeed, if (11) applies to all $\hat{\pi} \in (G_\tau(0), 1)$, then $G_\tau(\cdot)$ first-order stochastically dominates $G(\cdot)$ (see Dharmadhikari and Joag-Dev (1983)). More generally, as long as the transfer leads to (weakly) greater consumption by all types, $G_\tau(\cdot)$ first-order stochastically dominates $G(\cdot)$. Intuitively, when $q_\tau(\theta) \ge q(\theta)$ for all types, a given percentile in the distribution of quantities purchased after the transfer corresponds to a larger quantity than before the transfer. In this case, condition (11) simply amounts to a strengthening of the dominance ordering between $G_\tau(\cdot)$ and $G(\cdot)$ up to the $\pi_{\max}$th percentile.

By using (11), the properties that $F(\theta) = G(q)$ and $\theta'(q) = g(q)/f(\theta)$, and differentiating the local incentive compatibility condition $T'(q) = \theta(q)\nu'(q)$, it is possible to establish the next result.

COROLLARY 1: *Assume that there exists $M > 0$ such that $\nu'''(\cdot) \le M$ for M sufficiently small and that the transfer $\tau(\theta)$ leads to weakly greater consumption for all consumer types up to type $\theta_{\max}$ so that* (11) *holds. Then, $T_\tau''(q_\tau) \le T''(q)$ for all percentiles in the distributions of quantities purchased before and after the transfer up to the $\pi_{\max}$th percentile.*

When $T''(q) \le 0$, the transfer can then lead to greater price discrimination in the form of larger discounts.[23] In Section 5.4, we will show that this implication of our model is supported by the data.

4. IDENTIFICATION AND ESTIMATION

In this section, we discuss the identification and estimation of the model's primitives, which build on intuitions from Perrigne and Vuong (2010). Intuitively, the pricing behavior of a seller depends on the distribution of consumer types in a village market. Since this distribution can be mapped into that of purchased quantities, the distribution of consumer

[23]See the Supplemental Material 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 after the transfer.

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 these intuitions, here we derive estimators that rely on 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 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 that we document is likely to reflect only 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.

4.1. *Identification*

In a village market for a given good, 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 $[\underline{\theta}, \overline{\theta}]$, 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 utility 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 only of θ with $Y(\theta) \equiv I(\theta, q, \omega(\theta))$ and $q = q(\theta)$; see the discussion of the two-dimensional case in Appendix A.[24] Under standard assumptions, we show that these primitives are identified in each village from data on consumers' expenditures and purchases, which provide information about $T(q)$ and q, respectively. Note that $\overline{u}(\theta)$ and $Y(\theta)$ are identified only for households whose relevant constraints bind.[25] In light of the equivalence between the IR model and the BC model established above, 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 for each type, 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 a consumer of type θ who demands $\overline{q}(\theta)$. This approach is justified by the fact that the overwhelming majority of households 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 the quantities purchased of a good 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 nonparametrically identifiable for each good from information on prices (expenditures) and quantities purchased in our data, we treat them as known in our identification arguments.

[24] Marginal willingness to pay θ and absolute ability to pay w, as captured by household consumption and income, are highly correlated in our data, since the commodities we consider are normal goods.

[25] Any economy with reservation utility schedule $\overline{u}(\theta)$ or budget schedule $Y(\theta)$ binding on a subset of $[\underline{\theta}, \overline{\theta}]$ is observationally equivalent to an economy with the same primitives but reservation utility schedule $\widetilde{u}(\theta)$ or budget schedule $\widetilde{Y}(\theta)$ that agree with $\overline{u}(\theta)$ or $Y(\theta)$, respectively, on such a subset and are appropriately adjusted for the remaining types.

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 in (1) for each type. We use this latter condition to identify a seller's cost structure. However, by relying exclusively on information on prices and quantities, 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 aversion under the assumption that $v(\theta, q) = \theta\nu(q)$, which we maintain from now on. This specification of utility is ubiquitous in the literature on auctions and nonlinear pricing for its tractability (see Guerre, Perrigne, and Vuong (2000) and Perrigne and Vuong (2010)), so we consider it a natural benchmark.[26]

Marginal Cost and Multipliers on Constraints. The relationship between θ and q implied by incentive compatibility is central to the identification of the model—the monotonicity of $q(\theta)$ is hereafter maintained. 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 from (10) that since $q(\theta)$ is an increasing function, $F(\theta) = G(q)$ for $q = q(\theta)$ and so the cumulative distribution function of types is identified from that of quantities. The condition $F(\theta) = G(q)$ further implies that $f(\theta) = g(q)q'(\theta)$ for any $q = q(\theta)$. Given this mapping between the distributions 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

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

where $\varphi(q) \equiv d\log(\theta(q))/dq = \theta'(q)/\theta(q)$, and $c'(Q)$ replaces c. We next show 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 $\underline{q}$ to $\overline{q}$ yield 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 $\theta'(q)/\theta(q) = T''(q)/T'(q) + A(q)$ so that $\theta'(q)/\theta(q)$ is identified up to $A(q)$. Therefore, condition (12) also implies that

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

[26] 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, Guerre, Perrigne, and Vuong (2011). For some of the arguments here, we assume 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 Supplemental Material for this argument and a discussion of related results in the nonlinear pricing and hedonic pricing literatures.

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 $\varphi(q)$ is defined as $\theta'(q)/\theta(q)$, which is identified up to $A(q)$ as argued, 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 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.

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 from the value of $G(q)$ at the quantity at which $T'(q)$ equals $c'(Q)$ by (12).[27]

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

$$\begin{aligned} \log\bigl(\theta(q)\bigr) &= \log\bigl(\theta(\underline{q})\bigr) + \int_{\underline{q}}^{q} \frac{d\log\bigl(\theta(x)\bigr)}{dx}\, dx \\ &= \log\bigl(\theta(\underline{q})\bigr) + \int_{\underline{q}}^{q} \frac{g(x)\bigl[T'(x) - c'(Q)\bigr]}{T'(x)\bigl[\gamma\bigl(\theta(x)\bigr) - G(x)\bigr]}\, dx. \end{aligned} \tag{14}$$

Once $c'(Q)$ and $\gamma(\theta(q))$ are identified, $\theta(q)$ is also identified up to $\theta(\underline{q})$ by (14), since it is a known function of either identified or known objects.[28] Then, $f(\theta)$ is identified from $g(q)$ and the derivative $\theta'(q)$, since $f(\theta) = g(q)/\theta'(q)$ by $F(\theta) = G(q)$ at any $q = q(\theta)$, as argued.

PROPOSITION 6: *In a village, the support of consumers' marginal willingness to pay $\theta(q)$ is identified from the cumulative distribution function of quantities $G(q)$, the associated probability density function $g(q)$, the marginal price schedule $T'(q)$, the marginal cost of the total quantity provided $c'(Q)$, and the schedule of multipliers $\gamma(\theta(q))$ up to a level normalization. The probability density function of consumers' marginal willingness to pay $f(\theta)$ is identified from the probability density function of quantities $g(q)$ and the first derivative of $\theta(q)$.*

[27] When the standard model is *known* to apply, knowledge of $A(q)$ is unnecessary to identify $c'(Q)$ since $\gamma(\theta(q))$ equals 1 at all quantities, so $c'(Q)$ is identified from $T'(q)$ at the largest quantity.

[28] 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 at any quantity q^s such that $\gamma(\theta(q)) = G(q)$ and $T'(q) = c'(Q)$ if the slope of $\gamma(\theta(q))$ differs from $g(q)$ at such a quantity. Specifically, 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 $\gamma(\theta(q)) = G(q) + f(\theta(q))s_q(\theta(q), q)/v_{\theta q}(\theta(q), q)$.

Utility Function and Schedule of Reservation Utility. Note that knowledge of the coefficient of absolute risk aversion implies that the base marginal utility function $\nu'(q)$ is identified up to a level normalization. Once the marginal price schedule $T'(q)$ and the type support $\theta(q)$ are identified, though, $\nu'(q)$ is identified from them without the need for such a normalization by the incentive compatibility condition $\theta(q)\nu'(q) = T'(q)$. Then, we can recover $\nu(q)$ from $\nu'(q)$ up to its value at some quantity $q' = q(\theta')$ 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, the utility function $\theta(q)\nu(q)$ is identified. Moreover, $\overline{u}(\theta)$ is identified for all consumers whose participation (or budget) constraints bind ($d\gamma(\theta(q))/dq > 0$), 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 to a level normalization. The reservation utility (or budget) schedule is identified for all consumers whose participation (or budget) constraints bind.*

4.2. *Estimation*

We estimate the model separately in each village for each good 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 $\nu'(q)$. Specifically, the assumed expression for $T(q)$, the one for $G(q)$, and a seller's first-order condition for each quantity provide three estimating equations for the parameters of $T(q)$, $G(q)$, $\gamma(\theta(q))$, and for $c'(Q)$. Based on equation (14) and the local incentive compatibility condition $\nu'(q) = T'(q)/\theta(q)$, we estimate $\theta(q)$ and $\nu'(q)$ as known transformations of the parameterized functions $T'(q)$ (as implied by $T(q)$), $G(q)$, $\gamma(\theta(q))$, and of $c'(Q)$. In the second step, we estimate $f(\theta)$ from the estimated $\theta(q)$ via a kernel density estimator.[29]

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 be easily 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

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

where $T_{vj}(q_{vji}) \equiv E[p_{vj}(q_{vji})|q_{vji}]q_{vji}$, $p_{vj}(q_{vji})$ 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 $E[p_{vj}(q_{vji})|q_{vji}]$ 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. In particular, although multiple unit values may be associated with a same 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

[29]Note that if $f(\theta)$ is interpreted as the probability mass function associated with the empirical cumulative distribution function $G(q)$, this second step is unnecessary.

the quantities purchased and prices paid in a village provide direct information about the price schedule, which is a deterministic function of quantity according to 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 quantities purchased and $\Phi_{vj}(\cdot)$ is a flexible polynomial (up to the third degree, including a fractional polynomial). For each village and good, we select the specification of $\Phi_{vj}(\cdot)$ corresponding to the lowest value of the Akaike information criterion (AIC) for (16). Note that ε^{g}_{vji} in (16) captures not only recall or recording error but also error in observed purchase frequencies resulting from the timing of Progresa interviews. For instance, in the week preceding an 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 Multipliers on Constraints. 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

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

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 $x(\cdot)$ for good j in village v 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.[30] Then, expression (17) leads to

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

[30]Note that we have specified the multiplier as a smooth function of θ and so q. When constructing its predicted values, though, we allow the multiplier to increase discontinuously across quantities, and thus the relevant constraints to bind for subintervals of $[\underline{\theta}, \overline{\theta}]$, based on tests of the equality of the estimated values of the multiplier across consecutive quantities, as discussed below.

which is the sum of a linear-in-parameters function given by the first four terms, a non-linear one, and measurement error ε^s_{vji}, with $\underline{c}'_{vj}(Q_{vj}) \equiv c'_{vj}(Q_{vj})/t_{vj1}$, $\underline{\chi}_{vj0} \equiv \chi_{vj0}/t_{vj1}$, and $\underline{\chi}_{vj1} \equiv \chi_{vj1}/t_{vj1}$. For each village and good, we select the specifications of $x_{vj}(\cdot)$ and $\Gamma_{vj}(\cdot)$ associated with the lowest AIC value for (18).[31]

Support of Consumer Types and Base Marginal Utility Function. We normalize $\underline{\theta}$ to 1 and specify households' (log) marginal willingness to pay for good j in village v as

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

by (14), where N_{vj} is the number of households purchasing good j, Q_{vj} is the total quantity purchased in village v, $T'_{vj}(q_{vji})$ is derived from (15), and $\gamma_{vj}(q_{vji})$ and $G_{vj}(q_{vji})$ are specified as discussed.[32] Using the local incentive compatibility condition $\nu'(q) = T'(q)/\theta(q)$ and the form of $\theta(q)$, we estimate the base marginal utility from good j in village v as

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

Probability Density Function of Consumer Types. Given the estimated $\theta_{vji} = \theta_{vj}(q_{vji})$, we estimate the probability density function 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 model's fit 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 transfer on the unit prices of each good. We do so by exploiting the experimental variation in the data induced by the introduction of the transfer in a randomly selected subset of villages. Based on this reduced form, we also evaluate the ability of our model to account for the impact of Progresa on unit prices. See Appendix B and the Supplemental Material for omitted results and details, including the standard errors of the estimates of the model's parameters.

[31] To see how the parameters of (18) are identified, note that there exists at least one quantity q^{FB}_{vj} such that $\gamma_{vj}(q^{\mathrm{FB}}_{vj}) = G_{vj}(q^{\mathrm{FB}}_{vj})$, and the (mean) unit price of this quantity identifies $\underline{c}'_{vj}(Q_{vj})$. Since the multiplier at the largest quantity is equal to 1, differentiating (18) twice with respect to quantity and evaluating the resulting expressions at the largest quantity provide two conditions that pin down $\underline{\chi}_{vj0}$ and $\underline{\chi}_{vj1}$. 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 three additional quantities.

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

5.1. *Estimation Sample*

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

Sample Selection. We use five waves of the Progresa evaluation surveys, namely, October 1998, March 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 Supplemental Material, 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 in which 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.[33]

Unit Prices and Quantities Purchased. In the left panels of Figure 1, we report the schedule of mean unit prices by quantity purchased from each village by good computed as explained above (together with an interpolating solid line). Similarly, in the right panels, we report the corresponding cumulative distribution functions of quantities purchased. In most villages, the unit price of each good declines with quantity, which implies that unit prices are highest for the households that purchase the smallest quantities, and decreases more rapidly over the range of small quantities that most households purchase, as is evident by comparing the left and right 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 purchased, 0.1 kilos, is more than 8 pesos on average across villages, whereas the unit price of the largest quantity, 2 kilos, can be as low as 1.5 pesos.

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 model's fit 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.[34] 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 result from applying to the original samples of localities the same selection rules applied to the original samples of municipalities.

[33] In particular, focusing on villages with at least 50% of unit prices declining with quantity accounts for a small loss of villages, in which price schedules are markedly nonmonotone.

[34] In some villages, although $\log(\theta(q))$ is estimated, $\theta(q)$ is missing for some quantities when its value is exceedingly large, so $f(\theta)$ is not estimated for any such θ.

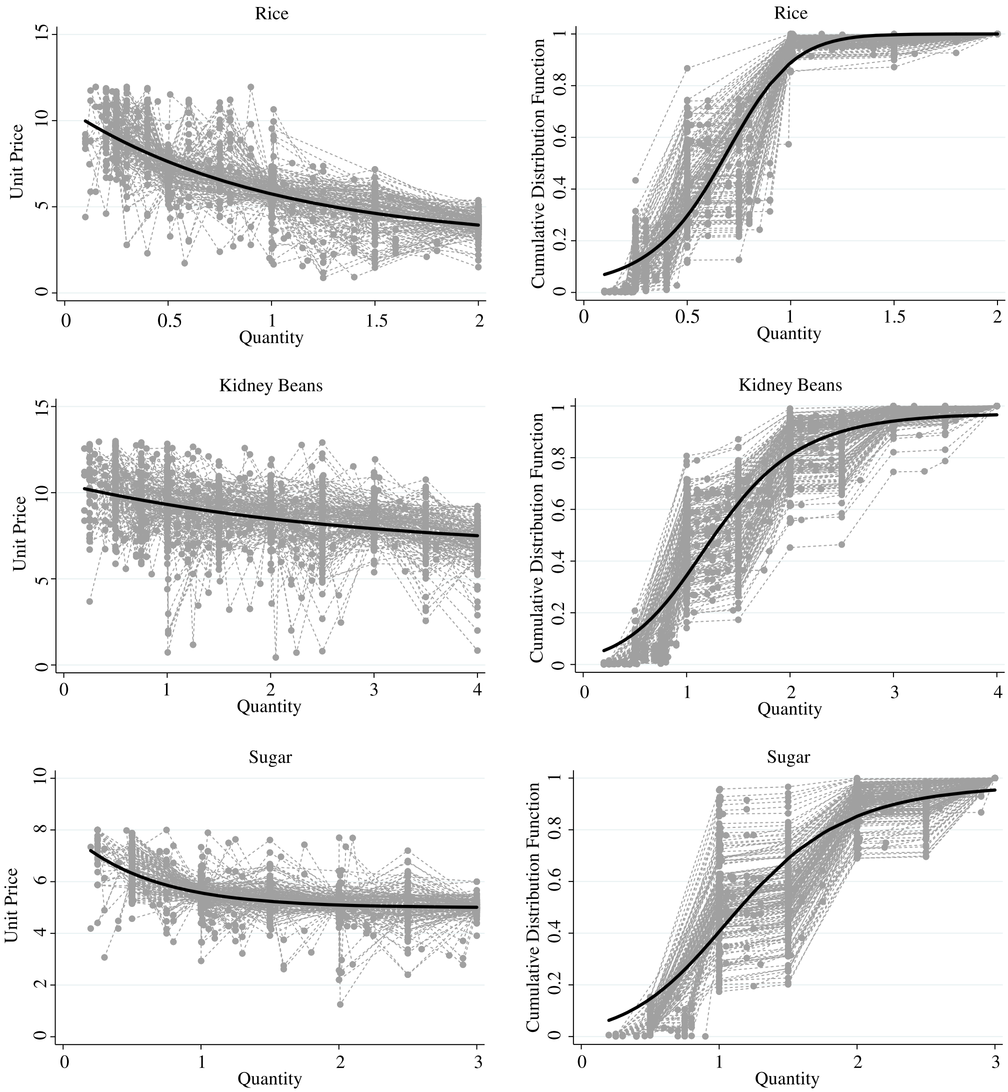

FIGURE 1.—Unit prices (left panels) and cumulative distribution functions of quantities (right panels).

Estimated Multipliers on Constraints. Figure 2 reports the estimated multiplier $\gamma(\theta(q))$ on participation (or budget) constraints for each quantity purchased from each village by good.[35] Recall that the multiplier ranges between 0 and 1 by construction. 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 1 for most of them: the 25th, 50th, and 75th percentiles in

[35]To obtain the estimated multiplier, we perform significance tests of the estimated values of $\gamma(\theta(q))$ in each village by good so as to determine whether the multiplier significantly differs across quantities. We then construct the predicted multiplier for each quantity accordingly. Note that if no parameter of $\gamma_{vj}(\cdot)$ is significant for good j in village v, then $\Gamma_{vj}(\cdot)$ equals 0 and the multiplier $\gamma_{vj}(\cdot)$ equals 0.5 at all quantities.

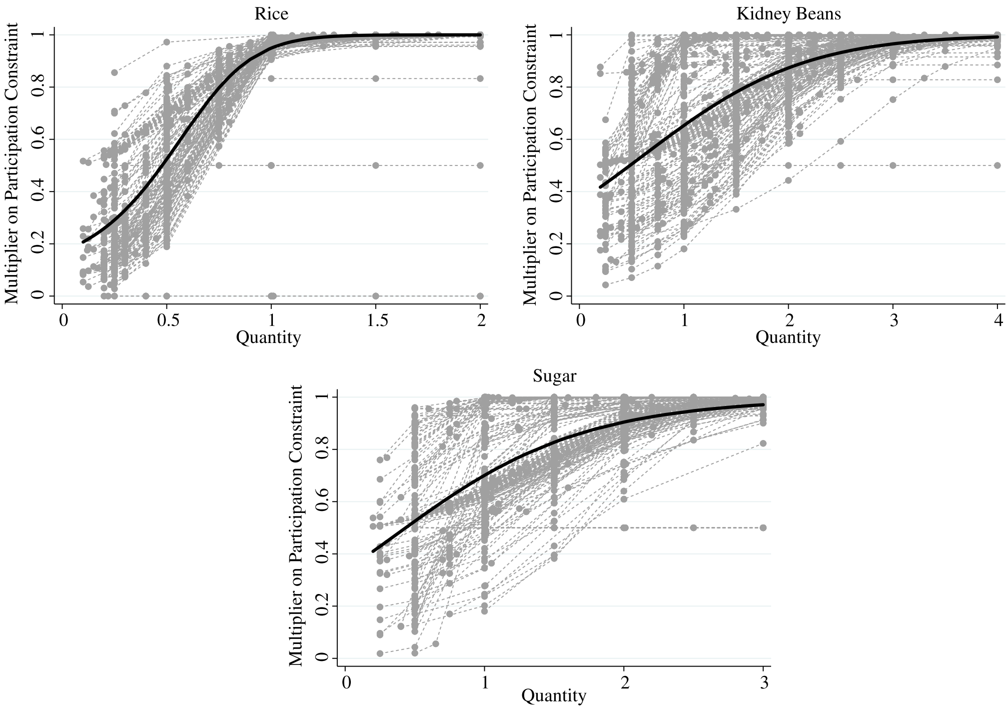

FIGURE 2.—Estimated multipliers on participation (or budget) constraints.

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.

As discussed, the shape of the multiplier function distinguishes different instances of our model. Note that the multiplier is estimated to be constant in only a handful of villages. Based on tests of the individual and joint significance of the estimated parameters of $\gamma(\theta(q))$, we reject the hypothesis that the standard model applies, that is, $\gamma(\theta(q)) = 1$ at all q, in nearly all villages. 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 the function $x(q)$ is positive and estimated to be roughly constant over the range of 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 the associated cumulative distribution function $G(q)$ starts at a low value and is weakly increasing. But an important departure from an 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.

Estimated Distributions of Consumer Types and Marginal Utility Functions. In the left panels of Figure 3, we report the estimates of base marginal utility, $\nu'(q)$, and in the right panels of the figure, we report the estimates of marginal utility, $\theta(q)\nu'(q)$, for each quantity purchased from each village by good. Note that $\nu'(q)$ decreases with the quantity

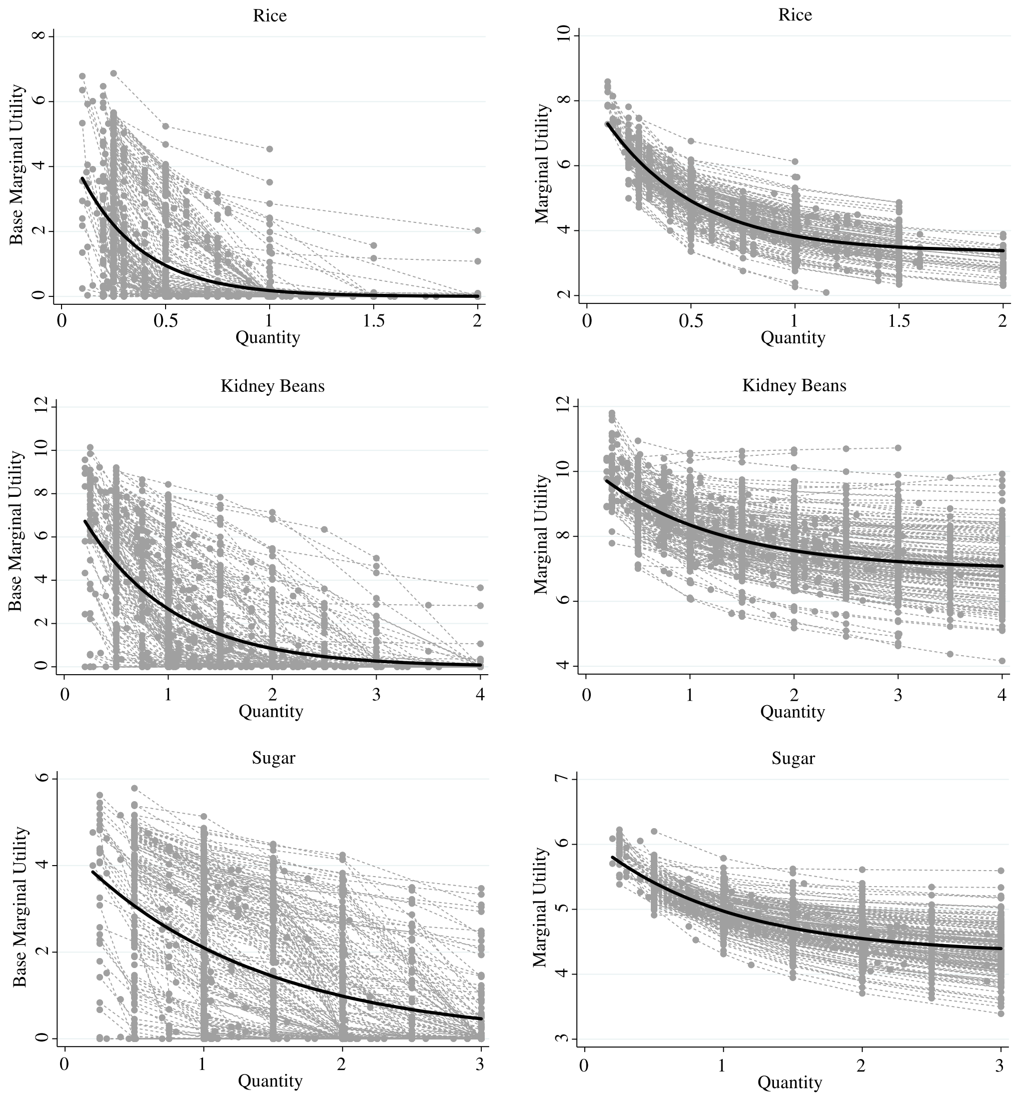

FIGURE 3.—Estimated base marginal utilities (left panels) and marginal utilities (right panels).

purchased in all villages, as is consistent with our model, although no such monotonicity restriction has been imposed in estimation. Instead, in nearly all villages, consumers' estimated marginal willingness to pay $\theta(q)$ increases with the quantity purchased, as is consistent with the incentive compatibility condition of our model, and rapidly so at large quantities. Because of this feature, the estimated support of consumer types for each good is much wider than that of quantities, as is evident from the distribution of (log) consumer types reported in Table II. Overall, marginal utility $\theta(q)\nu'(q)$ decreases with q for each good, although less rapidly than base marginal utility given that consumers' marginal willingness to pay $\theta(q)$ increases with q.[36]

[36]The estimated reverse hazard rate $f(\theta)/F(\theta)$ and hazard rate $f(\theta)/[1-F(\theta)]$ of the distribution of consumer types in each village for each good are mostly monotone with θ, respectively, weakly decreasing and

TABLE II

DISTRIBUTION OF LOG CONSUMER TYPES

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

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 taste and so in their marginal willingness to pay for a good, but also that different quantities of a good are valued quite differently by consumers of any given taste. We will explore these implications of nonlinear pricing in Section 5.3.

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 of a good in a village is determined by the cumulative distribution function of consumers' marginal willingness to pay $F(\theta)$, which satisfies $F(\theta) = G(q)$ at each $q = q(\theta)$; 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 ($c'(Q)$ in the general case). Although the distribution of marginal willingness to pay, the multiplier, and marginal cost are all unobserved, they are directly related to the observed distribution of quantities purchased and their unit prices by (17). Thus, one way to assess the model's fit 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 village and quantity purchased of the three goods, we plot in Figure 4 the estimated value of $G(q) - \gamma(\theta(q))$ on the y-axis against the estimated value of the markup measure $1/[t_1 p(q)] - 1/c'(Q)$ weighted by $x(q)$, or *weighted markup* for brevity, on the x-axis. A circle in any plot represents the model's fit to the data for a particular quantity of a good purchased in a village—the size of a circle reflects the fraction of households purchasing the quantity considered. By (17), then, the closer the relationship between $G(q) - \gamma(\theta(q))$ and the weighted markup to the 45-degree line, the better the model's fit to the data. Figure 4 shows that the model fits the price and quantity data well for each good. 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 first evaluate the degree of inefficiency of observed nonlinear pricing, and then assess its desirability by analyzing a counterfactual scenario in which sellers are prevented from discriminating and so price linearly.[37]

weakly increasing, which is one of the sufficient conditions for (PS); see the Supplemental Material. As neither of these restrictions has been imposed in estimation, we interpret these findings as validating our estimates of the type distributions.

[37]To reduce the impact of extreme observations, from this analysis of each good we exclude villages in which consumer types are estimated to be implausibly large, namely, in which the first quartile of the distribution of consumer types 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.

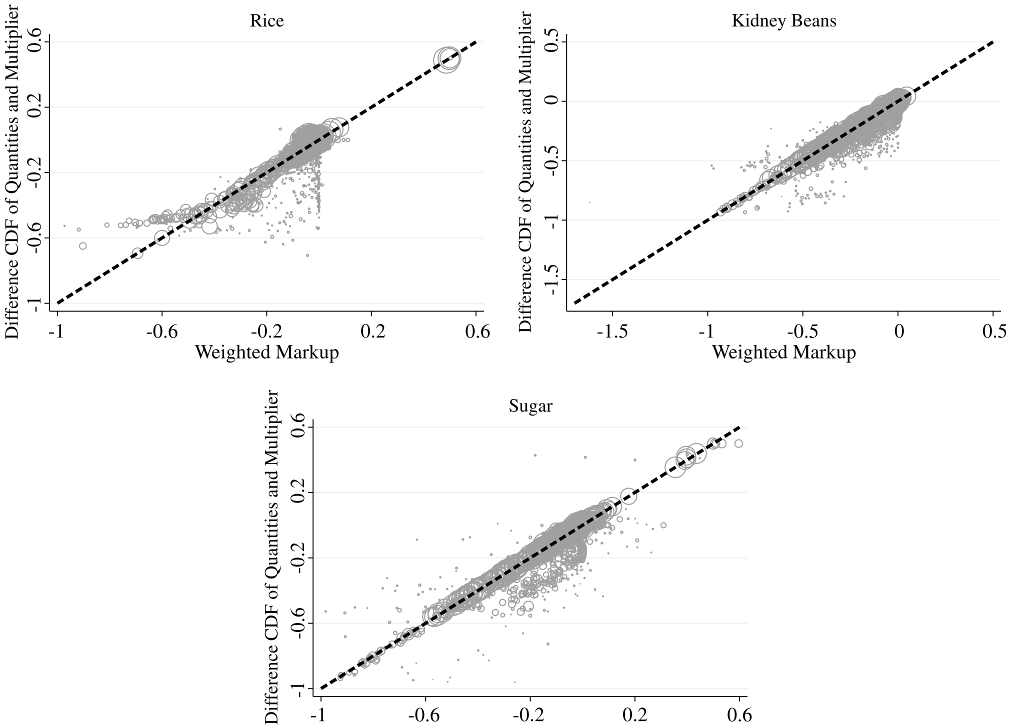

FIGURE 4.—Model fit within and across villages.

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 can distort the provision of a good relative to the first best, thereby reducing the gains from trade, and lead to allocations with levels of consumption below or *above* the first best. Accordingly, here we examine the distortions induced by sellers' market power for each good, as implied by our estimates, based on two measures: (i) the percentage difference between the 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 (17); and (ii) the fraction of households that purchase quantities larger than those under the first best, namely, with $\gamma(\theta(q)) < G(q)$.[38]

In the first five columns of Table III, we report the average percentage deviation of $\gamma(\theta(q))$ from $G(q)$ in absolute value ("Social Surplus Distortion") across villages by good and 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 that consume *above* the first best ("Overconsumption") across villages by good for the same percentiles in the distribution of consumer types in each village.

[38]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 $(x' + x)/2$.

TABLE III

NONLINEAR VERSUS PERFECTLY COMPETITIVE PRICING BY PERCENTILE RANGES OF CONSUMER TYPES

	Social Surplus Distortion					Overconsumption				
	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 III, the distortions associated with nonlinear pricing are larger for households that consume low to intermediate quantities in each village, that is, with low types. Households that 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, sellers overall practice an inefficient form of price discrimination, which, however, leads many 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 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 but also the *size* of the market served under the two pricing schemes: by Proposition 3, a seller who is prevented from discriminating may end up excluding some consumers. We find this to be the case for many villages in our sample. Namely, linear pricing leads to smaller consumer surplus and lower consumption for most consumers with low to intermediate valuations for kidney beans and sugar, including consumers of the smallest quantities. 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 of rice, kidney beans, and sugar tend to be better off under linear pricing.

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, $v'(q) = a[aq/(1-d) + b]^{d-1}$ with $a > 0$ and $aq/(1-d) + b > 0$.[39] To determine which households would participate under linear pricing, we

[39]Specifically, to limit parameter proliferation, we estimate the parameter d from a regression of estimated base marginal utility on quantity pooled across villages. 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. In this exercise, we interpret $c'(Q)$ as the marginal cost of a cost function with zero fixed costs and constant marginal cost.

TABLE IV
LINEAR PRICING (LP) VERSUS NONLINEAR PRICING (NLP) BY PERCENTILE RANGES OF CONSUMER TYPES

	Consumer Surplus under LP vs. NLP					Consumption under LP vs. NLP				
	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	54.7	17.9	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

also need an estimate of consumers' reservation utility. As discussed, though, consumers' reservation utility is identified only for 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 estimated type equal to this type's estimated utility under nonlinear pricing if the multiplier on this type's participation (or budget) constraint is significantly different from zero, which implies that the relevant constraint binds. We then specify any higher estimated type's reservation utility as equal to such type's estimated utility, if the multiplier on such type's constraint differs significantly from that on the next-lower estimated type's constraint, and equal to the next-lower estimated type's estimated utility otherwise.[40]

Based on these marginal utility and reservation utility schedules, in the first five columns of Table IV, we report the percentage of households across villages whose consumer surplus is higher under linear pricing than under nonlinear pricing by good 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 that consume more under linear pricing than under nonlinear pricing by good for the same percentiles in the distribution of consumer types in each village.

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 of the distribution of consumer types in each village would also consume *less* under linear pricing. Results for rice are different from those for kidney beans and sugar, since the estimated base marginal utility for rice is smaller over the range of quantities that most households buy, as apparent by comparing the right panels of Figure 1 and the left panels of Figure 3. Intuitively, the lower $v'(q)$, the lower the coefficient of absolute risk aversion $A(q)$ since $v'(q) = a^d A(q)^{1-d}$ by the assumed HARA form of $v(q)$ in this experiment, and so the higher $A(q)^{-1}$—we estimate that $d < 1$ for each good. In turn, a higher average $A(q)^{-1}$ implies a higher price elasticity of aggregate demand in absolute value under linear pricing, since this elasticity can be expressed as

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

[40]In the Supplemental Material, we consider the case in which the reservation utility is the maximal possible for each type, that is, $\overline{u}(\theta) = u(\theta)$, and obtain similar results.

and thus increases with $E_\theta[A(q_m(\theta))^{-1}]$; see the Supplemental Material for details. In fact, the median price elasticity of aggregate demand under linear pricing across villages is largest for rice, whereas the median marginal cost across villages is smallest for rice. Hence, unlike for kidney beans and sugar, in the case of rice, sellers do not have a strong incentive to charge high linear prices in the hope of attracting consumers with high valuations, who are willing to pay more. Indeed, both the mean and the median linear prices 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 induced by nonlinear pricing. We measure this effect by the percentage of households across villages that would not participate in the market under linear pricing. The percentages of excluded households in the percentile ranges of Table IV are 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 (in absolute value) price elasticity of aggregate demand under linear pricing 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 II, sellers more than make up for excluding low types by charging high prices to the remaining ones.

5.4. *The Effect of Income Transfers on Unit Prices*

In this subsection, we show that our model can account for a substantial fraction of the observed dispersion in the unit prices of rice, kidney beans, and sugar across quantities both within and across villages, as well as for the shift in the schedules of unit prices induced by Progresa and documented in Table I. Specifically, we first derive a reduced form of our model from a Taylor expansion of a seller's first-order condition, in which the slope of the unit price schedule of each good depends on the hazard rate of the distribution of quantities purchased in each village. We show that this reduced form explains a large fraction of the variation in unit prices across quantities and villages. We then show that once the dependence of unit prices on these hazard rates is explicitly accounted for, the direct effect of the program on the schedules of unit prices, in particular on their slopes, is insignificant. Hence, our model is sufficient to account for the impact of the program on the schedules of unit prices.

To elaborate, 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, Skoufias, and Washburn (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 schedules. In Table I, we have documented that Progresa has had a significant effect on unit prices in that it has led to an increase in the magnitude of quantity discounts. We have also shown that this effect cannot be detected without taking into account the nonlinearity of unit prices.

Our model is consistent with these findings. In terms of the insignificant impact of Progresa on average unit prices, the effect of an income transfer on unit prices, $p(q) = T(q)/q$, is ambiguous according to our model, since both households' expenditure, $T(q)$, and consumption, q, tend to increase. As for the importance of accounting for the nonlinearity of unit prices, although average unit prices may not increase after a transfer, the schedule of unit prices may nonetheless substantially change, as we established in Corollary 1, leading 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 the schedules of unit prices.

Transfers and Unit Prices. We examine the impact of the Progresa transfer on unit prices based on a second-order bivariate Taylor expansion of a seller's first-order condition in (9), interpreted as a function of $\log(q)$ and $[1 - G(q)]/g(q)$, in which $c'(Q)$ replaces c. Such an expansion yields

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

See Appendix A for details. In this expansion, the multiplier $\gamma(\theta(q))$ is interpreted as a function of quantity. This reduced form relates log (real) unit prices, $\log(p(q))$, to log quantities, $\log(q)$, and the inverse hazard rate of the distribution of quantities purchased, $[1 - G(q)]/g(q)$, 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-hand 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 quantities demanded, unit prices are then related to the hazard rate of the distribution of quantities purchased in a village, as (19) shows.

In Table I, we have documented a significant shift in the price schedules of the three goods of interest after the Progresa transfer. In Table V, we assess the extent to which our model can account for the observed nonlinearity of prices as well as for the change in price schedules resulting from Progresa. To this purpose, we first estimate (19) and then a version of it augmented to incorporate the impact of the Progresa transfer through both a "Treatment" dummy, which is equal to 1 for transactions occurring in localities targeted by the program, and an interaction term between this dummy and log quantity.[41] We stress that the inverse hazard rate of the distribution of quantities $[1 - G(q)]/g(q)$ in both regressions is computed for each locality. This approach thus allows for heterogeneous impacts of the program across localities, as captured by changes in the hazard rates

[41] Standard errors are computed by bootstrap at the locality level using 10,000 replication samples to account for the fact that hazard rates are estimated for each locality and wave. We could allow for locality fixed effects to capture the unobserved variability in marginal costs across localities—since we have several waves of data, both wave fixed effects and locality fixed effects are identified. The results we obtain by allowing for locality fixed effects are very similar to those in Table V. See the Supplemental Material, where we also report analogous results for alternative stratification and clustering schemes.

TABLE V

IMPACT OF CASH TRANSFERS ON PRICES BASED ON THE AUGMENTED MODEL (98% TRIMMING)[a]

	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

[a]Note: Wave fixed effects are included. Standard errors are clustered at the locality level.

of the distributions of quantities purchased, and could therefore be interpreted as a *mediation analysis* of the effect of Progresa on price schedules. To see that the program has affected the hazard rates of the distributions of quantities purchased in treated localities, note that the inverse hazard rate $[1 - G(q)]/g(q)$ for quantities in the top 25% of the distribution of unit prices in treated localities—namely, quantities with the highest unit prices that tend to be most nonlinear in quantity—on average is 18.6% higher for rice, 10.8% higher for kidney beans, and 17.6% higher for sugar relative to control localities. See the Supplemental Material for details.

Columns 1 of Table V report the estimates of (19). Observe that 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 term in the 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 across 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 term 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. In fact, the 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 I and are no longer significant for any good.

On the contrary, the interaction between $\log(q)$ and $[1 - G(q)]/g(q)$ in columns 2 of Table V is estimated to be significant for each good. Hence, in a precise sense, this statistic is sufficient to account for the impact of the program on unit prices: conditional on the interaction between $\log(q)$ and $[1 - G(q)]/g(q)$, the interaction between $\log(q)$ and "Treatment" no longer significantly affects price schedules. Specifically, the change in unit prices induced by the program is accounted for by the change in the distributions of quantities purchased, in particular by the change in their curvature as measured by the inverse hazard rate $[1 - G(q)]/g(q)$.

These results indicate that our model is capable of explaining the shift in price schedules reported in Table I. By the reduced form in (19), our model implies that unit prices depend on quantities purchased not just directly but also through the inverse hazard rate $[1 - G(q)]/g(q)$ of the distribution of quantities purchased in a village. We have shown that once the dependence of unit prices on $[1 - G(q)]/g(q)$ is taken into account, the effect of the program on the schedules of unit prices, in particular on their slopes, is no longer significant. Hence, the inverse hazard rate $[1 - G(q)]/g(q)$ captures the impact of the program.

Summary and Implications. The findings in Tables I and V support the key implication of our model that unit prices vary with quantity and that the relationship between unit prices and quantities purchased is affected not only by the distribution of consumer tastes but also by the distribution of consumer income. Based on the results in columns 2 of Table I, the Progresa transfer has not led to a significant change in average unit prices between treated and control localities. Unit prices, though, have changed substantially and differentially for consumers of small and large quantities in treated localities relative to control localities. For instance, the unit prices of the quantities in the *bottom* 25% of the distribution of quantities purchased across treated localities, paid by the households that purchase small quantities, on average are 13.2% *higher* for rice, 24.3% higher for kidney beans, and 29.8% higher for sugar than across control localities. On the contrary, the unit prices of the quantities in the *top* 25% of the distribution of quantities purchased across treated localities, paid by the households that purchase large quantities, on average are 12.3% *lower* for rice, 12.1% lower for kidney beans, and 5.6% lower for sugar than across control localities.[42] 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 a nonuniform price effect on both eligible and noneligible households. In light of the increase in the intensity of price discrimination that we document, then, the transfer may have had a more limited beneficial impact than has commonly been inferred.[43]

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 applicable to 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 implied by standard models of

[42]We have also estimated quantile treatment effects of the program on log unit prices and found significant changes between treated and control localities in line with the patterns discussed here. Results are similar when quantities and real expenditures are trimmed at the top 1% or 5%, rather than at the top 2% as in the sample of Table V. See the Supplemental Material for details.

[43]This argument, though, neglects the positive spillovers on noneligible households found by Angelucci and De Giorgi (2009).

nonlinear pricing, in which consumers are assumed to be unconstrained in their purchase decisions and outside options are presumed to 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 from a single market. We derive estimators of the model's primitives that can readily be implemented using a variety of publicly available data sets. 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. Specifically, we find that nonlinear pricing tends to lead to a greater degree of market participation, especially for consumers of small to intermediate quantities. This finding is all the more critical for the marginalized villages in our data, in which the consumption of several households is at subsistence levels.

We show that by increasing consumers' ability to pay, cash transfers provide sellers with an 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 prices 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 the unit prices of basic staples in our villages by inducing an overall shift in their schedules. Moreover, we 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 shift in price schedules. Namely, we document that the effect of the program on the schedules of unit prices is substantial once the dependence of unit prices on the quantities purchased is taken into account, although the program has not affected average unit prices. In particular, the program is associated with an increase in the intensity of price discrimination in the form of larger quantity discounts, which our model accounts for.

Our paper is one of the first to uncover changes in price schedules in villages included in the Progresa evaluation sample. This finding is especially relevant since cash transfers have become an increasingly popular poverty alleviation measure in Latin America and many other developing regions. Our 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.

APPENDIX A: OMITTED PROOFS AND DERIVATIONS

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). In analogy with assumption (PS) in the IR model, assumption (PS) in the BC model states that $l(\Phi, \theta)$ defined based on (7) weakly increases with θ for all $\Phi \in [0, 1]$. In the IR model, sufficient conditions for (PS) are

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

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 Supplemental Material. Recall that we consider the case in which the cost function $c(\cdot)$ is additively separable across consumers for simplicity. First, we rewrite the BC constraint as

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

since $u(\theta)=v(\theta,q(\theta))-t(\theta)$ by definition. For the moment, we presume that $\overline{u}$ is low enough that the (IR') constraints can be dropped. We show below that the (IR') constraints are satisfied under the conditions of Proposition 1. The BC problem can be expressed in Lagrangian-type form as

$$\begin{aligned}\max_{\{u(\theta)\},\{q(\theta)\}\in\widehat{Q}}&\left(\int_{\underline{\theta}}^{\overline{\theta}}\big[v\big(\theta,q(\theta)\big)-c\big(q(\theta)\big)-u(\theta)\big]f(\theta)\,d\theta\right.\\&\left.+\int_{\underline{\theta}}^{\overline{\theta}}\big\{I\big(\theta,q(\theta)\big)-\big[v\big(\theta,q(\theta)\big)-u(\theta)\big]\big\}\,d\Phi(\theta)\right) &(22)\\ \text{s.t.}\quad & u'(\theta)=v_\theta\big(\theta,q(\theta)\big), &(23)\end{aligned}$$

where $\widehat{Q}$ is the set of weakly increasing functions $q(\theta)$ and $\Phi(\theta)$ is the cumulative multiplier on the budget constraint expressed as in (21). Note that by adding and subtracting $\int_{\underline{\theta}}^{\overline{\theta}}u(\underline{\theta})f(\theta)\,d\theta$ to the term $\int_{\underline{\theta}}^{\overline{\theta}}u(\theta)f(\theta)\,d\theta$ in (22), we obtain

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

By using the local incentive compatibility condition $u'(\theta)=v_\theta(\theta,q(\theta))$ and integrating by parts, it follows

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

$$
- \int_{\underline{\theta}}^{\overline{\theta}} v_\theta\big(\theta, q(\theta)\big) F(\theta)\, d\theta
$$

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

Proceeding similarly, we can rewrite the term $\int_{\underline{\theta}}^{\overline{\theta}} u(\theta)\, d\Phi(\theta)$ in (22) as

$$
\begin{aligned}
&\int_{\underline{\theta}}^{\overline{\theta}} u(\theta)\, d\Phi(\theta) \\
&\quad = u(\underline{\theta})\big[\Phi(\overline{\theta}) - \Phi(\underline{\theta})\big] + \left(\int_{\underline{\theta}}^{\theta} v_\theta\big(x, q(x)\big)\, dx \right) \Phi(\theta) \Big|_{\underline{\theta}}^{\overline{\theta}} \\
&\qquad - \int_{\underline{\theta}}^{\overline{\theta}} v_\theta\big(\theta, q(\theta)\big) \Phi(\theta)\, d\theta \\
&\quad = u(\underline{\theta})\big[\Phi(\overline{\theta}) - \Phi(\underline{\theta})\big] + \Phi(\overline{\theta}) \int_{\underline{\theta}}^{\overline{\theta}} v_\theta\big(\theta, q(\theta)\big)\, d\theta \\
&\qquad - \int_{\underline{\theta}}^{\overline{\theta}} v_\theta\big(\theta, q(\theta)\big) \Phi(\theta)\, d\theta.
\end{aligned} \tag{25}
$$

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

$$
\begin{aligned}
&\int_{\underline{\theta}}^{\overline{\theta}} \big[v\big(\theta, q(\theta)\big) - c\big(q(\theta)\big)\big] f(\theta)\, d\theta + \int_{\underline{\theta}}^{\overline{\theta}} \big[I\big(\theta, q(\theta)\big) - v\big(\theta, q(\theta)\big)\big]\, d\Phi(\theta) - u(\underline{\theta}) \\
&\quad - \int_{\underline{\theta}}^{\overline{\theta}} v_\theta\big(\theta, q(\theta)\big)\, d\theta \\
&\quad + \int_{\underline{\theta}}^{\overline{\theta}} v_\theta\big(\theta, q(\theta)\big) F(\theta)\, d\theta + u(\underline{\theta})\big[\Phi(\overline{\theta}) - \Phi(\underline{\theta})\big] + \Phi(\overline{\theta}) \int_{\underline{\theta}}^{\overline{\theta}} v_\theta\big(\theta, q(\theta)\big)\, d\theta \\
&\quad - \int_{\underline{\theta}}^{\overline{\theta}} v_\theta\big(\theta, q(\theta)\big) \Phi(\theta)\, d\theta,
\end{aligned}
$$

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

$$
\begin{aligned}
&\int_{\underline{\theta}}^{\overline{\theta}} \big[v\big(\theta, q(\theta)\big) - c\big(q(\theta)\big)\big] f(\theta)\, d\theta + \int_{\underline{\theta}}^{\overline{\theta}} \left[\frac{F(\theta) - \Phi(\theta) + \Phi(\overline{\theta}) - 1}{f(\theta)} \right] v_\theta\big(\theta, q(\theta)\big) f(\theta)\, d\theta \\
&\quad + \int_{\underline{\theta}}^{\overline{\theta}} \frac{\phi(\theta)\big[I\big(\theta, q(\theta)\big) - v\big(\theta, q(\theta)\big)\big]}{f(\theta)} f(\theta)\, d\theta + u(\underline{\theta})\big[\Phi(\overline{\theta}) - \Phi(\underline{\theta}) - 1\big].
\end{aligned} \tag{26}
$$

By an argument similar to the one in the proof of Result 1 in the Supplemental Material, it is possible to show that $\Phi(\overline{\theta}) = 1$. Then, by collecting terms one more time and dropping irrelevant constants, it is immediate that the expression in (26) reduces to the objective function in (6). The following result is analogous to Result 4 in the Supplemental Material.

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 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 $u_{\mathrm{IR}}(\theta)$, $q_{\mathrm{IR}}(\theta)$, $t_{\mathrm{IR}}(\theta)$, $\overline{u}_{\mathrm{IR}}(\theta)$, and $\overline{q}_{\mathrm{IR}}(\theta)$. To start, by Result 4 in the Supplemental Material, 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

$$v_q\big(\theta, q(\theta)\big) - c'\big(q(\theta)\big) = \left[\frac{\gamma(\theta) - F(\theta)}{f(\theta)}\right] v_{\theta q}\big(\theta, q(\theta)\big) \tag{27}$$

for each type and the complementary slackness condition

$$\int_{\underline{\theta}}^{\overline{\theta}} \big[u(\theta) - \overline{u}(\theta)\big]\, d\gamma(\theta) = 0 \tag{28}$$

hold. By Result 4 in the Supplemental Material, the optimal allocation in the IR problem is unique.

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

$$\begin{aligned} v_q\big(\theta, q(\theta)\big) - c'\big(q(\theta)\big) = {} & \left[\frac{\Phi(\theta) - F(\theta)}{f(\theta)}\right] v_{\theta q}\big(\theta, q(\theta)\big) \\ & + \frac{\phi(\theta)\big[v_q\big(\theta, q(\theta)\big) - I_q\big(\theta, q(\theta)\big)\big]}{f(\theta)} \end{aligned} \tag{29}$$

for each type and the complementary slackness condition

$$\int_{\underline{\theta}}^{\overline{\theta}} \big[I\big(\theta, q(\theta)\big) - v\big(\theta, q(\theta)\big) + u(\theta)\big]\, d\Phi(\theta) = 0 \tag{30}$$

hold, together with $u(\theta) \geq \overline{u}$. Note that for $\Phi(\theta)$ to be a legitimate cumulative multiplier, it must be nonnegative and weakly increasing with θ. Let $\Phi(\theta) = \gamma(\theta)$ 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, (30), 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 (29) reduce to those of the IR problem in (27). Hence, the quantity profile that solves the IR problem satisfies the first-order conditions of the BC problem. In the third step, we show that given this quantity profile, consumers reach the same utility in the two problems.

Step 1: Verify That the Complementary Slackness Condition of the BC Problem Holds, $t(\theta) \le I(\theta, q(\theta))$*, and* $u(\theta) \ge \overline{u}$. We first claim that (30) holds at the IR allocation and the IR allocation satisfies $t(\theta) \le I(\theta, q(\theta))$. To this purpose, recall that, by assumption, $I(\theta, q(\theta)) \ge v(\theta, q(\theta)) - \overline{u}(\theta)$, $I(\theta, q(\theta)) = v(\theta, q(\theta)) - \overline{u}(\theta)$ for types whose IR constraints bind, and $\overline{u}(\underline{\theta}) \ge \overline{u}$. Note that when the IR constraints bind in that $d\gamma(\theta) = d\Phi(\theta) > 0$, then $q(\theta) = \overline{q}(\theta)$, $u(\theta) = v(\theta, q(\theta)) - t(\theta) = \overline{u}(\theta)$, and thus $t(\theta) = v(\theta, q(\theta)) - \overline{u}(\theta)$. 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 in that $d\gamma(\theta) = d\Phi(\theta) = 0$, then $u(\theta) = v(\theta, q(\theta)) - t(\theta) \ge \overline{u}(\theta)$ and so $t(\theta) \le v(\theta, q(\theta)) - \overline{u}(\theta)$. Given that, by assumption, $v(\theta, q(\theta)) - \overline{u}(\theta) \le I(\theta, q(\theta))$ for types whose IR constraints do not bind, it follows that $t(\theta) \le I(\theta, q(\theta))$. Hence, if condition (28) holds for the IR problem, then condition (30) holds for the BC problem. Also, $t(\theta) \le I(\theta, q(\theta))$ is satisfied, and $u(\theta) \ge \overline{u}$ holds by (IR) and the fact that $\overline{u}(\underline{\theta}) \ge \overline{u}$ by assumption.

Step 2: Verify That the First-Order Conditions of the BC Problem Reduce to Those of the IR 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-hand side of (29) equals zero for each θ, and so the first-order conditions of the BC problem in (29) are identical to those of the IR problem in (27).

Step 3: Verify That the IR and BC Problems Imply the Same Utility. We argue that 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 and Step 2 ensure that the utility achieved by each consumer from the quantity profile $\{q(\theta)\}$ is identical in the IR and BC problems. Specifically, consider a type θ' whose IR constraint binds. Since $u'(\theta) = v_\theta(\theta, q(\theta))$ by local incentive compatibility and, by assumption, $u(\theta') = \overline{u}(\theta') = v(\theta', \overline{q}(\theta')) - I(\theta', \overline{q}(\theta'))$, the utility of any type θ higher than θ' in the IR problem is

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

The right-hand side of the second equality in (31) is the utility that the consumer achieves in 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. An analogous argument holds for any type lower than θ'.

This establishes the desired result. *Q.E.D.*

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 (29) and the complementary slackness condition (30) are satisfied with $I(\theta, q(\theta, w), w) = Y(w) - \underline{z}(\theta, q(\theta, w))$. Our next result shows how this menu varies across submarkets. For this, let

$$\begin{aligned} &t(\theta, w_H) = t(\theta, w_L) + Y(w_H) - Y(w_L), \quad q(\theta, w_H) = q(\theta, w_L), \quad \text{and} \\ &\Phi(\theta, w_H) = \Phi(\theta, w_L). \end{aligned} \tag{32}$$

RESULT 2: *If the implementable allocation $\{u(\theta, w_L), q(\theta, w_L)\}$ with associated cumulative multipliers $\{\Phi(\theta, w_L)\}$ solves the simple BC problem in submarket w_L, then the implementable allocation $\{u(\theta, w_H), q(\theta, w_H)\}$ with associated cumulative multipliers $\{\Phi(\theta, w_H)\}$ satisfying* (32) *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 menu 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,

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

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 (29) and the complementary slackness condition (30) with $u(\theta)$, $t(\theta)$, $q(\theta)$, $\Phi(\theta)$, $\phi(\theta)$, and $I(\theta, q(\theta))$ replaced by $u(\theta, w_L)$, $t(\theta, w_L)$, $q(\theta, w_L)$, $\Phi(\theta, w_L)$, $\phi(\theta, w_L)$, and $I(\theta, q(\theta, w_L), w_L)$. It is immediate that the allocations and multipliers implied by (32) satisfy the corresponding first-order and complementary slackness conditions for submarket w_H. To see why, note that since $I_q(\theta, q, w) = -\underline{z}_q(\theta, q)$, the first-order conditions in the two submarkets are identical under (32). Consider next the complementary slackness condition. Since this condition holds for 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

$$t(\theta, w_L) = I\big(\theta, q(\theta, w_L), w_L\big) = Y(w_L) - \underline{z}\big(\theta, q(\theta, w_L)\big). \tag{34}$$

But then the multiplier $\phi(\theta, w_H)$ in submarket w_H for a consumer with taste θ is also positive under (32), since

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

where the first and third equalities follow from (32) and the second equality follows from (34). An analogous argument applies when the budget constraint of a consumer with taste θ in submarket w_L does not bind. Hence, the conjectured allocation satisfies the first-order and complementary slackness conditions for submarket w_H. So, by Result 1, the conjectured allocation solves the simple BC problem for submarket w_H. *Q.E.D.*

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 weakly monotone function of θ, namely, $w = \omega(\theta)$ with $\omega'(\theta) \geq 0$. Substituting $w = \omega(\theta)$ into $I(\theta, q, w) = Y(w) - \underline{z}(\theta, q)$ gives

$$I\big(\theta, q, \omega(\theta)\big) = Y\big(\omega(\theta)\big) - \underline{z}(\theta, q) \tag{35}$$

for any q. Under (35), the analogues of Result 1 and Proposition 1 apply. *Q.E.D.*

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

$$T''(q) = \theta'(q)\nu'(q) + \theta(q)\nu''(q) = \theta(q)\nu'(q)\left[\frac{\theta'(q)}{\theta(q)} + \frac{\nu''(q)}{\nu'(q)}\right]$$
$$= T'(q)\left[\frac{1}{\theta(q)q'(\theta)} - A(q)\right]. \tag{36}$$

Observe that by using the fact that $T'(q) = \theta(q)\nu'(q)$, 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

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

Note that $\theta > [\gamma(\theta) - F(\theta)]/f(\theta)$ by the seller's first-order condition since $\nu'(\cdot), c > 0$. Hence, the denominator of $q'(\theta)$ is positive, and so is the numerator of $q'(\theta)$ since $q'(\theta) > 0$ is a maintained assumption. By substituting the above expression for $q'(\theta)$ into (36) and using the fact that $T'(q), A(q) > 0$, we can express $T''(q) \leq 0$ as

$$T'\big(q(\theta)\big)A\big(q(\theta)\big)\left\{\frac{\theta - \dfrac{\gamma(\theta) - F(\theta)}{f(\theta)}}{\theta\dfrac{d}{d\theta}\left[\theta - \dfrac{\gamma(\theta) - F(\theta)}{f(\theta)}\right]} - 1\right\} \leq 0 \quad \Leftrightarrow$$

$$\frac{\theta - \dfrac{\gamma(\theta) - F(\theta)}{f(\theta)}}{\theta \dfrac{d}{d\theta}\left[\theta - \dfrac{\gamma(\theta) - F(\theta)}{f(\theta)}\right]} \leq 1. \tag{37}$$

Recall that the IR constraints can bind at isolated points only for $\underline{\theta}$ or $\overline{\theta}$. Hence, we can partition $(\underline{\theta}, \overline{\theta})$ into (sub)intervals of types whose IR constraints bind and (sub)intervals of types whose IR constraints do not bind. Then, there are only two cases to consider, depending on whether a consumer belongs to an interval of types whose IR constraints bind or an interval of types whose IR constraints do not bind.

Consider first an interval of types whose IR constraints bind. By construction, any such type θ purchases $\overline{q}(\theta)$ and achieves utility $\overline{u}(\theta)$. Assumption (H) implies that $\overline{u}'(\theta) = \nu(\overline{q}(\theta))$ so that $\overline{q}'(\theta) = \overline{u}''(\theta)/\nu'(\overline{q}(\theta))$. Then, by (36),

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

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 an interval of types whose IR constraints do not bind, in which case $\gamma(\theta) = \gamma$ for any such type. When $\gamma = 1$, the above expression for $q'(\theta)$ becomes

$$q'(\theta) = \frac{\dfrac{d}{d\theta}\left[\theta - \dfrac{1 - F(\theta)}{f(\theta)}\right]}{\left[\theta - \dfrac{1 - F(\theta)}{f(\theta)}\right] A\big(q(\theta)\big)} \geq \frac{1}{\theta A\big(q(\theta)\big)} = \frac{\nu'\big(q(\theta)\big)}{-\theta\nu''\big(q(\theta)\big)}, \tag{38}$$

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 (38) implies that $1/q'(\theta) \leq -\theta\nu''(q(\theta))/\nu'(q(\theta))$, which combined with (36) yields that

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

When, instead, $\gamma \in [0, 1)$, the last inequality in (37) becomes

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

We prove that (39) holds by considering two further cases in which $\gamma < 1$.

Case 1: $\gamma \geq F(\theta)$. In this case, $[\gamma - F(\theta)]f(\theta) \geq 0$ so that a sufficient condition for (39) is

$$f^2(\theta) \geq -\big[\gamma - F(\theta)\big]f'(\theta). \tag{40}$$

If $f'(\theta) \geq 0$, then it is immediate that (40) is satisfied. Suppose now that $f'(\theta) < 0$. Recall that $d([1 - F(\theta)]/f(\theta))/d\theta \leq 0$ by assumption. Thus,

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

which is equivalent to $f^2(\theta) \geq -[1 - F(\theta)]f'(\theta)$. Moreover, $-[1 - F(\theta)]f'(\theta) > -[\gamma - F(\theta)]f'(\theta)$ since $f'(\theta) < 0$ and, by the assumption of the case, $\gamma < 1$. Then, it follows that (40) and so (39) are satisfied.

Case 2: $\gamma < F(\theta)$. Note that we can rewrite (39) as

$$\theta f(\theta) \geq F(\theta) - \gamma + \left[F(\theta) - \gamma\right]\frac{\theta f'(\theta)}{f(\theta)}. \tag{41}$$

When $f'(\theta) \leq 0$, a sufficient condition for (41) is $F(\theta) \leq \theta f(\theta)$, which is assumed. Thus, (39) holds. When, instead, $f'(\theta) > 0$, a sufficient condition for (41) is $\theta f(\theta) \geq F(\theta) + F(\theta)\theta f'(\theta)/f(\theta)$, which can be expressed as

$$f^2(\theta) - F(\theta)f'(\theta) \geq \frac{f(\theta)}{f(\theta)} \cdot \frac{F(\theta)f(\theta)}{\theta} = \frac{f^2(\theta)F(\theta)}{\theta f(\theta)}$$

or, equivalently, as

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

which holds by assumption. Hence, (39) is satisfied. *Q.E.D.*

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 weakly 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

$$u(\theta) = \theta\nu\big(q(\theta)\big) - T\big(q(\theta)\big) > u_m(\theta) = \theta\nu\big(q_m(\theta)\big) - \theta\nu'\big(q_m(\theta)\big)q_m(\theta), \tag{42}$$

where in (42) 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

$$\theta\nu\big(q_m(\theta)\big) - \theta\nu'\big(q_m(\theta)\big)q_m(\theta) \geq \theta\nu\big(q(\theta)\big) - \theta\nu'\big(q_m(\theta)\big)q(\theta), \tag{43}$$

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 (42) and (43) imply that $\theta\nu(q(\theta)) - T(q(\theta)) > \theta\nu(q(\theta)) - \theta\nu'(q_m(\theta))q(\theta)$, that is,

$$\theta\nu'\big(q_m(\theta)\big) > T\big(q(\theta)\big)/q(\theta) = p\big(q(\theta)\big). \tag{44}$$

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

$$T'\big(q(\theta)\big) \leq T\big(q(\theta)\big)/q(\theta) = p\big(q(\theta)\big). \tag{45}$$

This inequality, together with (44) and the incentive compatibility condition $T'(q(\theta)) = \theta\nu'(q(\theta))$, implies

$$\theta\nu'\big(q(\theta)\big) = T'\big(q(\theta)\big) \leq T\big(q(\theta)\big)/q(\theta) < \theta\nu'\big(q_m(\theta)\big), \tag{46}$$

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 $q \geq q(\theta)$, $q(\theta) \geq q_m(\theta)$, and $\gamma(\theta) < 1$, then the utility of a consumer of type θ is weakly higher under linear pricing than under nonlinear pricing. Consider one such type, say, $\hat{\theta}$. By way of contradiction, suppose that $u(\hat{\theta}) > u_m(\hat{\theta})$. We will show that if this is the case, 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 exists if $\gamma(\hat{\theta}) < 1$. To reach the desired contradiction, rewrite $u(\hat{\theta}) > u_m(\hat{\theta})$ as

$$u(\theta_2) - \big[u(\theta_2) - u(\hat{\theta})\big] > u_m(\theta_2) - \big[u_m(\theta_2) - u_m(\hat{\theta})\big], \tag{47}$$

which can equivalently be expressed as

$$u(\theta_2) - \int_{\hat{\theta}}^{\theta_2} u'(x)\,dx > u_m(\theta_2) - \int_{\hat{\theta}}^{\theta_2} u_m'(x)\,dx. \tag{48}$$

Condition (48), in turn, is equivalent to

$$\overline{u}(\theta_2) - u_m(\theta_2) > \int_{\hat{\theta}}^{\theta_2} \big[\nu\big(q(x)\big) - \nu\big(q_m(x)\big)\big]\,dx \tag{49}$$

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-hand side of (49) is nonnegative, which establishes the desired contradiction. To see that the right-hand side of (49) is nonnegative, note first that for all $\theta \geq \hat{\theta}$,

$$p_m = \theta\nu'\big(q_m(\theta)\big) = \hat{\theta}\nu'\big(q_m(\hat{\theta})\big) \geq \hat{\theta}\nu'\big(q(\hat{\theta})\big) = T'\big(q(\hat{\theta})\big) \geq T'\big(q(\theta)\big) = \theta\nu'\big(q(\theta)\big), \tag{50}$$

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 $q(\hat{\theta}) \geq q_m(\hat{\theta})$ 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 $q \geq q(\hat{\theta})$ by assumption and $q(\theta) \geq q(\hat{\theta})$ for $\theta \geq \hat{\theta}$. Hence, (50) 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 $q(\theta) \geq q_m(\theta)$ for all $\theta \geq \hat{\theta}$ and $\nu(\cdot)$ increasing imply that the right-hand side of (49) is nonnegative, which, in turn, yields 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, let $\hat{\theta}$ be a type in $[\theta', \theta'']$ with $u_m(\hat{\theta}) = \overline{u}(\hat{\theta})$. To establish the desired claim, we need to show that at least one consumer type in $(\hat{\theta}, \theta'']$ does not participate under linear pricing, although such a consumer type participates under nonlinear pricing by assumption (FP). To this purpose, suppose, by way of contradiction, that all types in $(\hat{\theta}, \theta'']$ 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 $\theta \in (\hat{\theta}, \theta'']$ who participates under linear pricing, it must be that $u_m(\theta) \geq \overline{u}(\theta)$, which can be expanded as

$$u_m(\theta) = u_m(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} u_m'(x)\,dx = u_m(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} \nu\big(q_m(x)\big)\,dx$$
$$\geq \overline{u}(\theta) = \overline{u}(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} \overline{u}'(x)\,dx = \overline{u}(\hat{\theta}) + \int_{\hat{\theta}}^{\theta} \nu\big(\overline{q}(x)\big)\,dx, \tag{51}$$

where the second equality in (51) uses the fact that $u_m'(\theta) = \nu(q_m(\theta))$ by the definition of $u_m(\theta)$ and the consumer's first-order condition $\theta\nu'(q_m(\theta)) = p_m$ under linear pricing, 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, (51) implies

$$\int_{\hat{\theta}}^{\theta} \nu\big(q_m(x)\big)\,dx \geq \int_{\hat{\theta}}^{\theta} \nu\big(\overline{q}(x)\big)\,dx. \tag{52}$$

With $\nu(\cdot)$ positive and increasing, (52) implies that there exists a subinterval of types between type $\hat{\theta}$ and type $\theta \leq \theta''$ with positive measure, say, $[\theta_1, \theta_2]$, such that $q_m(\theta''') \geq \overline{q}(\theta''')$ for all $\theta''' \in [\theta_1, \theta_2]$. Since $\overline{q}(\theta) > q^{\mathrm{FB}}(\theta)$ by assumption for all $\theta \in [\theta', \theta'']$ and $q_m(\theta''') \geq \overline{q}(\theta''')$ for all $\theta''' \in [\theta_1, \theta_2]$, it follows that $q_m(\theta''') > q^{\mathrm{FB}}(\theta''')$ for all $\theta''' \in [\theta_1, \theta_2]$. Combining $q_m(\theta''') > q^{\mathrm{FB}}(\theta''')$ for all $\theta''' \in [\theta_1, \theta_2]$ with the fact that $\nu'(\cdot)$ is decreasing gives that

$$p_m = \theta'''\nu'\big(q_m(\theta''')\big) < \theta'''\nu'\big(q^{\mathrm{FB}}(\theta''')\big) = c \tag{53}$$

for any such θ''', 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''', $q^{\mathrm{FB}}(\theta''')$. But the condition $p_m < c$ implied by (53) contradicts seller optimality under linear pricing. *Q.E.D.*

PROOF OF PROPOSITION 4: By the equivalence between the IR model and the BC model, consider the BC model. Recall that $\{t(\theta), q(\theta)\}$ denotes the optimal menu before the transfer is introduced with $q(\theta) = \overline{q}(\theta)$ for consumer types who spend all of their budgets on the seller's good. Denote by $\{t_\tau(\theta), q_\tau(\theta)\}$ the optimal menu after the transfer is introduced with $q_\tau(\theta) = \overline{q}_\tau(\theta)$ for consumer types who spend all of their budgets on the seller's good. Before the transfer, the budget for the seller's good is $I(\theta, \overline{q}(\theta)) = Y - \underline{z}(\theta, \overline{q}(\theta))$ for quantity $\overline{q}(\theta)$. By (BCH), $\overline{t}(\theta) = I(\theta, \overline{q}(\theta))$ so that

$$\overline{t}'(\theta) = I_\theta\big(\theta, \overline{q}(\theta)\big) + I_q\big(\theta, \overline{q}(\theta)\big)\overline{q}'(\theta). \tag{54}$$

After the transfer, the budget for the seller's good is $I(\theta, \overline{q}_\tau(\theta), \tau(\theta)) \equiv Y + \tau(\theta) - \underline{z}(\theta, \overline{q}_\tau(\theta))$ for quantity $\overline{q}_\tau(\theta)$, which can be expressed as

$$I\big(\theta, \overline{q}_\tau(\theta), \tau(\theta)\big) = I\big(\theta, \overline{q}_\tau(\theta)\big) + \tau(\theta). \tag{55}$$

By (BCH) after the transfer, $\bar{t}_\tau(\theta) = I(\theta, \overline{q}_\tau(\theta), \tau(\theta))$. This latter condition and (55) yield that

$$\bar{t}'_\tau(\theta) = I_\theta\big(\theta, \overline{q}_\tau(\theta)\big) + I_q\big(\theta, \overline{q}_\tau(\theta)\big)\overline{q}'_\tau(\theta) + \tau'(\theta). \tag{56}$$

Recall that $\{\overline{q}_{\text{IR}}(\theta)\}$ denotes the reservation quantity profile in the IR model. By Proposition 1, $\overline{q}_{\text{IR}}(\theta) = \overline{q}(\theta)$ for types with binding BC constraints in the BC model before the transfer. This fact, the condition $\theta\nu'(\overline{q}(\theta))\overline{q}'(\theta) = \bar{t}'(\theta)$ by (BCH), (54), and the requirement of Proposition 1 that $\theta\nu'(\overline{q}_{\text{IR}}(\theta)) = I_q(\theta, \overline{q}_{\text{IR}}(\theta))$ imply that $I_\theta(\theta, \overline{q}(\theta)) = 0$ before the transfer. Similarly, after the transfer, it must be that $I_\theta(\theta, \overline{q}_\tau(\theta)) + \tau'(\theta) = 0$. Hence, $I_\theta(\theta, \overline{q}_\tau(\theta)) = -\tau'(\theta) > 0$ and so $I_\theta(\theta, \overline{q}_\tau(\theta)) > I_\theta(\theta, \overline{q}(\theta))$ for any consumer type with transfer that satisfies $\tau'(\theta) < 0$. It follows that $\overline{q}_\tau(\theta) > \overline{q}(\theta)$ for any such consumer type since $I_{\theta q}(\cdot, \cdot) \geq 0$.

We now argue that there exists an interval of types with $q_\tau(\theta) > q(\theta)$ and $T'_\tau(q_\tau(\theta)) < T'(q(\theta))$. By assumption, there exists a consumer of type θ' whose budget constraint binds before and after the transfer with $q_\tau(\theta') = \overline{q}_\tau(\theta') > \overline{q}(\theta') = q(\theta')$ by the argument in the previous paragraph. From local incentive compatibility and the fact that $\nu''(\cdot) < 0$, it follows that $T'_\tau(q_\tau(\theta')) = \theta'\nu'(q_\tau(\theta')) < \theta'\nu'(q(\theta')) = T'(q(\theta'))$ for type θ'. Note that, by continuity, $q_\tau(\theta) > q(\theta)$ for an interval of types containing θ'. Hence, again from local incentive compatibility and the fact that $\nu''(\cdot) < 0$, it follows that $T'_\tau(q_\tau(\theta)) = \theta\nu'(q_\tau(\theta)) < \theta\nu'(q(\theta)) = T'(q(\theta))$ for types in such an interval. Thus, there exists an interval of types Θ' containing θ' such that $q_\tau(\theta) > q(\theta)$ and $T'_\tau(q_\tau(\theta)) < T'(q(\theta))$.

As shown, the transfer amounts to an expansion in consumers' budgets for the seller's good. In particular, $T_\tau(q_\tau(\theta')) > T(q(\theta'))$ for type θ' given that $\tau(\theta') > 0$, $q_\tau(\theta') > q(\theta')$, and $\underline{z}(\theta, q)$ weakly decreases with q. Hence, the price schedule must increase for at least a subinterval of Θ'. *Q.E.D.*

PROOF OF COROLLARY 1: By local incentive compatibility before and after the transfer, we have that $T'(q) = \theta(q)\nu'(q)$ and $T'_\tau(q_\tau) = \theta_\tau(q_\tau)\nu'(q_\tau)$, where $T(q)$ and $T_\tau(q_\tau)$ 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 both local incentive compatibility conditions gives

$$T''_\tau\big(q_\tau(\theta)\big) = \frac{\nu'\big(q_\tau(\theta)\big)}{q'_\tau(\theta)} + \theta\nu''\big(q_\tau(\theta)\big) \quad \text{and} \quad T''\big(q(\theta)\big) = \frac{\nu'\big(q(\theta)\big)}{q'(\theta)} + \theta\nu''\big(q(\theta)\big). \tag{57}$$

Recall from (10) that $F(\theta) = G(q)$ for $q = q(\theta)$ and, similarly, $F(\theta) = G_\tau(q_\tau)$ for $q_\tau = q_\tau(\theta)$. Therefore, differentiating $F(\theta) = G(q(\theta))$ and $F(\theta) = G_\tau(q_\tau(\theta))$ gives $f(\theta) = g(q(\theta))q'(\theta) = g_\tau(q_\tau(\theta))q'_\tau(\theta)$, which in turn yields that $1/q'_\tau(\theta) \leq 1/q'(\theta)$ if, and only if, $g_\tau(q_\tau(\theta)) \leq g(q(\theta))$, since

$$\frac{1}{q'_\tau(\theta)} = \frac{g_\tau(q_\tau(\theta))}{f(\theta)} \leq \frac{g(q(\theta))}{f(\theta)} = \frac{1}{q'(\theta)}. \tag{58}$$

Observe that (11) states that $g_\tau(q_\tau(\theta)) \leq g(q(\theta))$ up to the $\pi_{\max}$th percentile in the distributions of quantities purchased before and after the transfer, where $q_\tau(\theta) = G_\tau^{-1}(\hat{\pi})$

and $q(\theta) = G^{-1}(\hat{\pi})$ for $\hat{\pi} \in (G_\tau(0), \pi_{\max})$ by definition of $q_\tau(\theta)$ and $q(\theta)$. Hence, $q'(\theta) \leq q'_\tau(\theta)$ by (58) up to some type $\theta_{\max}$ such that $\hat{\pi}_{\max} = G(q(\theta_{\max})) = G_\tau(q_\tau(\theta_{\max}))$. The assumptions that $q_\tau(\theta) \geq q(\theta)$ for $\theta \leq \theta_{\max}$ and $\nu''(\cdot) \leq 0$ further imply that $\nu'(q_\tau(\theta)) \leq \nu'(q(\theta))$ for $\theta \leq \theta_{\max}$. Moreover, since, by assumption, $\nu'''(\cdot) \leq M$ and M can be chosen small enough, the difference between $\nu''(q_\tau(\theta))$ and $\nu''(q(\theta))$ can be made as small as desired for all $\theta \leq \theta_{\max}$. Thus, it is immediate by (57) that $T''_\tau(q_\tau(\theta)) \leq T''(q(\theta))$ for all percentiles in the distributions of quantities purchased before and after the transfer up to the $\pi_{\max}$th percentile, which corresponds to quantity $q(\theta_{\max})$ before the transfer and quantity $q_\tau(\theta_{\max})$ after the transfer. *Q.E.D.*

PROOF OF PROPOSITION 5: Rewrite the seller's first-order condition in (9) as

$$\begin{aligned} \frac{1}{T'(q)} &= \frac{1}{c'(Q)} + \frac{[F(\theta) - \gamma(\theta)]}{c'(Q)\theta f(\theta)} = \frac{1}{c'(Q)} + \frac{[G(q) - \gamma(\theta(q))]\theta'(q)}{c'(Q)g(q)\theta(q)} \\ &= \frac{1}{c'(Q)} + \frac{[G(q) - \gamma(\theta(q))]\varphi(q)}{c'(Q)g(q)}, \end{aligned} \tag{59}$$

where $q = q(\theta)$, $c'(Q)$ replaces c, $g(q) = f(\theta)\theta'(q)$, and $\varphi(q) = d\log(\theta(q))/dq = \theta'(q)/\theta(q)$. Equivalently,

$$\frac{g(q)}{\varphi(q)}\left[\frac{c'(Q)}{T'(q)} - 1\right] = G(q) - \psi(q), \tag{60}$$

where $\psi(q) \equiv \gamma(\theta(q))$. By further differentiating (60), we obtain

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

Integrating this expression from $\underline{q}$ to $\overline{q}$ gives

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

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 that

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

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. *Q.E.D.*

DERIVATION OF THE REDUCED FORM IN (19): With $c'(Q)$ replacing c, the seller's first-order condition in (9) can be rewritten as

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

since $F(\theta)=G(q)$, $f(\theta)=g(q)q'(\theta)$, and $q'(\theta)=1/\theta'(q)$, or, equivalently, as

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

Hence,

$$\log\left(\frac{c'(Q)}{T'(q)}\right)\approx\frac{\theta'(q)}{\theta(q)}\left[\frac{G(q)}{g(q)}-\frac{\gamma\big(\theta(q)\big)}{g(q)}\right].$$

Further manipulating this expression and using the fact that $T'(q)=t_1 p(q)$ by our specification for $T(q)$ in (15) yield that

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

which implies that

$$\log\big(p(q)\big)\approx\log\left(\frac{c'(Q)}{t_1}\right)-\frac{\theta'(q)}{\theta(q)}\left[\frac{1-\gamma\big(\theta(q)\big)}{g(q)}\right]+\frac{\theta'(q)}{\theta(q)}\left[\frac{1-G(q)}{g(q)}\right].\tag{61}$$

Letting $\psi(q)\equiv\gamma(\theta(q))$, (61) can be expressed as

$$\log\big(p(q)\big)\approx\log\left(\frac{c'(Q)}{t_1}\right)-\frac{\theta'\big(e^{\log(q)}\big)}{\theta\big(e^{\log(q)}\big)}\left[\frac{1-\psi\big(e^{\log(q)}\big)}{g\big(e^{\log(q)}\big)}\right]+\frac{\theta'\big(e^{\log(q)}\big)}{\theta\big(e^{\log(q)}\big)}\left[\frac{1-G(q)}{g(q)}\right].\tag{62}$$

Then, we can interpret the right-hand side of (62) 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

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

Therefore,

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

$$+ \underbrace{\left[f_1(a,b) - af_{11}(a,b) - bf_{12}(a,b)\right]}_{\beta_1} \log(q)$$

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

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

which can equivalently be expressed as

$$\log\bigl(p(q)\bigr) \approx \beta_0 + \beta_1 \log(q) + \beta_2 \left[\frac{1 - G(q)}{g(q)}\right] + \beta_3 \log(q) \left[\frac{1 - G(q)}{g(q)}\right] + \beta_4 [\log(q)]^2 + \beta_5 \left[\frac{1 - G(q)}{g(q)}\right]^2. \qquad \text{Q.E.D.}$$

APPENDIX B: OMITTED ESTIMATION RESULTS

We present here estimation results omitted from the main text.

Estimated Marginal Costs. Figure 5 reports the estimated marginal cost of the total quantity provided of each good in each village (municipality). The mean estimated marginal cost across villages 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. Since the villages we study are fairly dispersed and isolated, some variability in estimated marginal cost across villages is to be expected. The mean and range of the estimated marginal cost of each good across villages are nonetheless very similar across goods.

Statistics on Estimates. The first three columns of Tables VI to VIII report the quartiles of the distributions 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 these estimates. Since these estimates, except for $c'(Q)$, vary across quantities in each village, the last three columns of these tables report the quartiles of the distribution across villages of the village-level median t-statistic of each of these estimates. These statistics are meant to illustrate the variability of the precision of these estimates across villages. As apparent from these tables, the model's primitives as well as the multiplier $\gamma(\theta(q))$ are mostly significantly different from zero.

Finally, omitting village and good subscripts, Table IX reports the quartiles of the distributions across villages of the t-statistics of the estimates of the parameters of equations (15), (16), and (18) from each village except for $c'(Q)$, whose estimates are reported in Tables VI to VIII. Recall that equations (15), (16), and (18) correspond, respectively, to the estimated specifications for $T(q)$, $G(q)$, and a seller's first-order condition for each village and good. The parameters reported in Table IX are: (i) (t_0, t_1) for equation (15); (ii) $(\phi_0, \phi_1, \phi_2, \phi_3, \phi_4, \phi_5)$ for equation (16) depending on the chosen specification for $\Phi(q)$, namely, (ϕ_0, ϕ_1) when $\Phi(q)$ is a polynomial of degree 1, (ϕ_0, ϕ_1, ϕ_2) when $\Phi(q)$

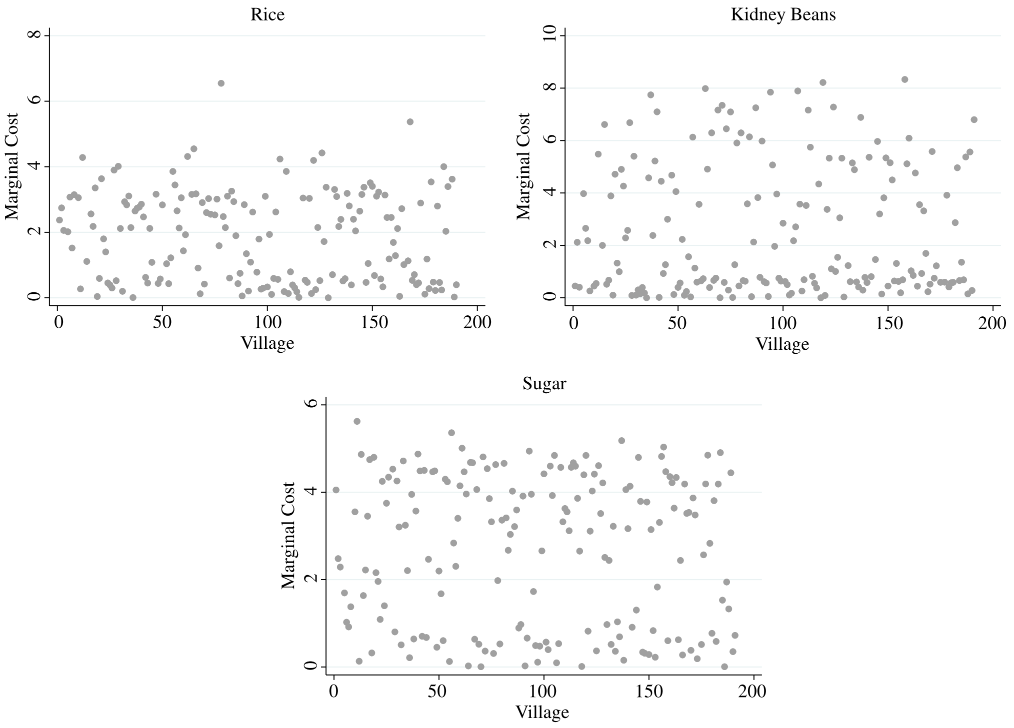

FIGURE 5.—Estimated marginal costs.

TABLE VI

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 VII

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 VIII

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 IX

DISTRIBUTION OF t-STATISTICS OF PARAMETER ESTIMATES 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
ϕ_0	14.884	30.924	52.712	2.199	3.755	6.112	3.662	7.533	12.846
ϕ_1	13.273	30.277	52.464	0.397	1.503	3.502	1.221	3.031	20.417
ϕ_2	4.759	15.067	34.319	1.221	1.709	2.677	1.706	2.839	3.802
ϕ_3	0.943	4.525	19.341	13.920	13.920	13.920	35.639	35.639	35.639
ϕ_4	12.631	26.762	45.275	1.991	3.116	4.545	7.015	9.717	15.506
ϕ_5	8.477	17.251	35.139	6.000	13.116	20.780	29.268	40.037	62.714
$\underline{\chi}_0$	5.245	11.914	37.571	2.290	4.422	14.124	3.270	8.169	30.075
$\underline{\chi}_1$	2.784	8.973	37.825	4.606	8.923	29.260	6.206	14.297	38.293
γ_0	13.815	32.662	106.822	6.216	15.209	41.150	6.615	19.618	70.537
γ_1	7.431	19.279	1187.317	2.099	8.861	52.094	8.461	15.658	43.505
γ_2	17.552	49.238	101.920	4.453	11.238	30.227	5.446	12.633	26.724

is a polynomial of degree 2, $(\phi_0, \phi_1, \phi_2, \phi_3)$ when $\Phi(q)$ is a polynomial of degree 3, or (ϕ_0, ϕ_4, ϕ_5) when $\Phi(q)$ is the fractional polynomial $\Phi(q) = \phi_0 + \phi_4 q^{\phi_5}$; (iii) $\underline{\chi}_1$ or $(\underline{\chi}_0, \underline{\chi}_1)$ for the auxiliary function $x(q)$ in equation (18) depending on the chosen specification for $x(q)$; and (iv) $(\gamma_0, \gamma_1, \gamma_2)$ for the index $\Gamma(q)$ of the multiplier function in equation (18) depending on the chosen degree of the polynomial specification for $\Gamma(q)$, namely, $(\gamma_0, \gamma_1, \gamma_2)$, (γ_1, γ_2), (γ_0, γ_2), (γ_0, γ_1), γ_2, γ_1, or none, in which case $\Gamma(q) = 0$ and the multiplier equals 0.5.[44] As apparent from the table, these parameters are also for the most part significantly different from zero. See the Supplemental Material for the confidence intervals of all these estimated parameters including $c'(Q)$.

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. [208,242,245]

ATTANASIO, O., AND C. FRAYNE (2006): "Do the Poor Pay More?" Unpublished Manuscript. [208,213,240]

[44]The parameter t_1 is constrained so as to ensure that the estimated $T(q)$ is an increasing and concave function of q, whereas the parameters of $\Phi(q)$ and $\Gamma(q)$ are constrained so as to guarantee that both functions weakly increase with quantity. We also impose an upper bound on $c'(Q)$ consistent with a seller making nonnegative profits.

ATTANASIO, O., AND E. PASTORINO (2020): "Supplement to 'Nonlinear Pricing in Village Economies'," *Econometrica Supplemental Material*, 88, https://doi.org/10.3982/ECTA13918. [214]
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. [212,226]
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. [229]
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. [208,216]
CUNHA, J., G. DE GIORGI, AND S. JAYACHANDRAN (2017): "The Price Effects of Cash versus In-Kind Transfers." Unpublished Manuscript. [211]
DEATON, A. (1989): "Rice Prices and Income Distribution in Thailand: A Non-Parametric Analysis," *Economic Journal*, 99 (395), 1–37. [213]
DHARMADHIKARI, S. W., AND K. JOAG-DEV (1983): "Mean, Median, Mode III," *Statistica Neerlandica*, 37 (4), 165–168. [227]
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. [207]
GUERRE, E., I. PERRIGNE, AND Q. VUONG (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68 (3), 525–574. [229]
HODDINOTT, J., E. SKOUFIAS, AND R. WASHBURN (2000): "The Impact of Progresa on Consumption: A Final Report," Washington, DC: International Food Policy Research Institute. [208,214,242]
JENSEN, R. T., AND N. H. MILLER (2008): "Giffen Behavior and Subsistence Consumption," *American Economic Review*, 98 (4), 1553–1577. [219,220]
JULLIEN, B. (2000): "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93 (1), 1–47. [207,209,215-218,222,225,247]
LANCASTER, K. J. (1966): "A New Approach to Consumer Theory," *Journal of Political Economy*, 74 (2), 132–157. [220]
MASKIN, E., AND J. RILEY (1984): "Monopoly With Incomplete Information," *RAND Journal of Economics*, 15 (2), 171–196. [207,208,215]
MCINTOSH, A. C. (2003): "The Poor Pay More for Their Water," *Habitat Debate*, 9 (3), 12. [207]
PANNARUNOTHAI, S., AND A. MILLS (1997): "The Poor Pay More: Health-Related Inequality in Thailand," *Social Science and Medicine*, 44 (12), 1781–1790. [207]
PERRIGNE, I., AND Q. VUONG (2010): "Nonlinear Pricing in Yellow Pages." Unpublished Manuscript. [227, 229]
STOLE, L. A. (2007): "Price Discrimination and Competition," in *Handbook of Industrial Organization*, Vol. 3, ed. by M. Armstrong and R. Porter, Elsevier. [219]
TADELIS, S., AND I. SEGAL (2005): "Lectures in Contract Theory." Unpublished Manuscript. [216]
THOMAS, L. (2002): "Non-Linear Pricing With Budget Constraint," *Economics Letters*, 75 (2), 257–263. [216]

The Econometric Society
Cowles Foundation
Yale University
30 Hillhouse Avenue
New Haven, CT 06511, USA
http://www.econometricsociety.org