The Causal Effect of Competition on Prices and Quality: Evidence from a Field Experiment[1]

Matias Busso
Research Department
Inter-American
Development Bank

Sebastian Galiani
Department of Economics
University of Maryland

April 2014

Abstract

This paper provides the first experimental evidence on the effect of increased competition on the prices and quality of goods. We rely on an intervention that randomized the entry of 61 retail firms (grocery stores) into 72 local markets in the context of a conditional cash transfer program that serves the poor in the Dominican Republic. Six months after the intervention, product prices in the treated districts had decreased by about 6%, while product quality and service quality had not changed. Our results are also informative to the design of social policies. They suggest that policymakers should pay attention to supply conditions even when the policies in question will only affect the demand side of the market.

JEL: D4, L1, I3.

Keywords: Competition, prices, quality, experimental evidence, design of social policies.

[1] We thank Francisca Müller and Juan Sebastian Muñoz for their excellent project management and research assistance. We thank very useful comments by Daron Acemoglu, Gary Libecap, James Poterba, Ernesto Schargrodsky, Gustavo Torrens, and seminar participants at various places. This experiment has been registered in the American Economic Association RCT Registry with number AEARCTR-0000343. The views expressed herein are those of the authors and should not be attributed to the Inter-American Development Bank.

1. Introduction

Ever since Adam Smith, economists have seen market competition as a way of achieving economic efficiency. If a competitive equilibrium exists, then the equilibrium is necessarily Pareto optimal in the sense that there is no other allocation of resources which would make all participants in the market better off. Adam Smith considered competition to be a form of rivalry between suppliers that eliminated excessive profits, removed excessive supplies and satisfied existing demand (Stigler, 1957). Competition also exerts downward pressure on costs, reduces slack periods and provides incentives for the efficient organization of production (Nickell, 1996). Price-taking implies that no supplier is able to exert market power, which means that firms do not price profitably above the marginal cost of production and that consumer surplus is therefore maximized. All these sound arguments notwithstanding, real-world experimental evidence on the welfare effects of competition has not, to the best of our knowledge, been presented.

In this paper we exploit a randomized control experiment to assess what impact the entry of new grocery stores into a market has on prices and quality. The experiment was part of an attempt to improve the operations of the Dominican Republic conditional cash transfer (*Solidaridad*) program, which provides monetary transfers to poor families that can only be spent using a debit card that is accepted only by a network of grocery stores that are affiliated with the program. In those stores, program recipients can use the money to purchase authorized goods (basically, products belonging to a specified set of food items). Under the original program design, the retail stores in the program network wielded market power, and the government argued that they were using this power to raise prices and to offer a more limited variety of products than the product range offered by stores outside the network. This was seen as signaling a loss of consumer surplus and therefore a potential welfare loss. In response to this situation, we collaborated with the Dominican Republic government to evaluate the effects of the expansion of the retail network with the goal of encouraging competition among stores.

The intervention was conducted during May and June 2011 and involved bringing 61 new grocery stores into the network in 72 districts. The experimental design allowed anywhere from zero to three stores to start operating in each district.

We use data collected at baseline and six months after the intervention on both retailers and households located in the areas concerned. Our data allow us to arrive at precise price measurements and then to use these prices to infer quality. The surveys also incorporate other independent measures of quality. We estimate average treatment effects using the randomization assignment in order to instrument the potentially endogenous entry of new stores that is induced by noncompliance with randomization.

We find that entry into the market leads to a significant and robust reduction in prices but that it did not lead to any change in the quality of the goods or the quality of the delivery of those goods by the grocery stores. We also impose some structure in order to estimate the price-elasticity at entry, which we find to be 0.08.

Previous work has relied on observational data. Trapani and Oslon (1982) analyze the effect of the deregulation of the airline industry in the US on the price and quality of service. This paper analyzes the relationship between fare level, open entry and quality of service. It exploits a cross-sectional sample of 70 city-pair markets within the United States in 1971 and 1977. The authors found that increasing competition in the airline industry leads to a reduction both in prices and in the average quality of service. Their paper shows that the independent effect of decreasing market concentration, which leads to a higher quality of service, is overshadowed by the independent effect of price competition (lower prices), which, in turn, lowers the quality of service. Bresnahan and Reiss (1990) empirically model the effects of entry into monopolistic and duopolistic markets. They study the interrelationships among potential entrants' profit levels and decisions. Using cross-sectional data on 149 geographically isolated US markets for new automobiles, they estimate that the second entrant has nearly the same costs and market opportunities as the first entrant. They also find that entry does not cause price-cost margins to fall by a large margin.

Several papers have developed detail econometric models of the competitive effects of market entry, including those of Carlton (1983), Berry (1989), Bresnashan and Reiss (1989) and Reiss and Spiller (1989). Bresnahan and Reis (1991) take this area of research further by proposing an empirical framework for measuring the effects of entry into concentrated markets. They develop a model of entry in which firms have U-shaped average costs and entrants face entry barriers, and they then use this model to estimate entry thresholds ratios that provide a measure of the market size required to support a given number of firms. These estimations were obtained from cross-sectional data on the number of firms in 202 distinct geographic markets, with the sample being limited to five retail and professional service industries. The authors rely on ordered probit models to estimate the entry thresholds and the equilibrium number of market entrants. Their empirical results suggest that competitive behavior changes quickly as market size and the number of incumbents increase. In markets with five or fewer incumbents, almost all variation in competitive behavior occurs with the entry of the second and third firms, at which point prices fall and then level off.

Geroski (1989) examines a dynamic feedback model of entry and profit margins applied to panel data covering a six-year period (1974-1979) for 85 UK three-digit industries. He finds that entry barriers are rather high in most industries and that there are noticeable differences in the pace of competitive dynamics. The author concludes that entry penetration leads to a reduction in prices because competition from entrants may trigger a decrease in X-inefficiency and units costs.

Besker and Noel (2009) analyze the effect of Wal-Mart's entry into the grocery market using a store-level price panel dataset. They find that competitors' response to the entry of a Wal-Mart Supercenter, which has a price advantage over competitors of about 10% percent, is a price reduction of 1%-1.2%, on average, with most of this reduction being accounted for by smaller-scale competitors. They conclude that competitors' responses vary in line with their degree of differentiation from Wal-Mart. At one extreme, the largest supermarket chains reduce their prices by less than half as much as its smaller competitors. At the opposite extreme, low-end grocery stores, which compete more directly with Wal-Mart, cut their prices more than twice as much as higher-end stores. Jia (2008) develops an empirical model --one which relaxes the assumption that entry into different markets is independent-- to assess the impact of Kmart stores on Wal-Mart stores and other discount retailers and to quantify the size of the scale economies obtained within a given chain. She finds that the negative impact of Kmart's presence on Wal-Mart's profits was much stronger in 1988 than in 1997, while the opposite is true for the effect of Wal-Mart's presence on Kmart's profits.

In a more recent paper, Bennett and Yin (2013) explore the relationship between market development and drug quality by evaluating the impact of chain-store entry (Med-Plus) into the Indian pharmaceutical industry. They rely on a difference-in-differences identification strategy and find that the entry of a chain store leads to a relative 5% improvement in quality, measured on the basis of compliance with the standards of the Indian Pharmacopeia Commission, and a 2% decrease in price. The authors conclude that the chain store appears to increase retail competition by offering higher-quality drugs and lower prices. Although this evidence is very interesting, the effects associated with chain-store entry cannot be unequivocally attributed to an increase in competition, since the new stores operate on the basis of a completely different rationale than the incumbent family firms do. As the authors argue, it is better to interpret the evidence they have gathered as being indicative of the effects of market development in a developing country.[2]

[2] There has been a great deal of cross-sectional econometric analysis of the relationship between market structure and innovative activity. More specifically, Geroski (1990) finds that concentration and other measures of market power tend to be associated with lower rates of innovation. By the same token, Porter (1990) finds that competition leads to a higher rate of innovation and that it is a significant factor in generating world-beating industries. In addition, in recent years there have been a number of comprehensive studies on the relationship between technical efficiency and market structure; the findings of these studies are reported in Caves (1990), Green and Mayes (1991) and in Caves and associates (1992). In these papers, the central finding is that an increase in market concentration above a certain threshold tends to reduce firm efficiency. On the other hand, organization theory puts some stress on the notion of a match between the structure of a firm and the environment in which it operates. On the same subject, Donaldson (1987) finds that a more competitive environment generates the incentive for a rapid adjustment of the structure in the direction of a good fit, while Caves (1980) shows that competition prompts companies to employ more efficient decision-making structures. Relying on panel data for British manufacturing firms, Blundell et al. (1999), controlling for firm heterogeneity, find a robust and positive effect of market share on innovation. Aghion et

Although several important studies have used different setups to focus on the effects of the entry of new competitors into imperfectly competitive markets, our paper contributes to the literature by reporting, to the best of our knowledge, on the first field randomized-control experiment undertaken to assess the impact of increasing competition on prices and quality. This is significant because, as has been acknowledged in the literature, competition in observational data studies is likely to be endogenous for the parameters of interest (see, among others, Blundell et al. (1999) and Aghion et al. (2014)).

Finally, in a study that is somewhat related to the subject of interest here although it deals with a different topic, Hasting and Washington (2008) suggest that grocery stores face a large and predictable increase in demand for the goods that are most heavily purchased by beneficiaries of the Food Stamps Program. They use two years of item-level scanner data from three Nevada stores belonging to a national supermarket chain to examine the stores´ pricing response to this demand shift. They find that the increase in aggregate demand induced by benefit delivery results in small food-price increases. Related to this work is the paper by Cunha et al. (2013), which examines the effect of cash versus in-kind transfers on local prices in Mexico. Both types of transfers increase the demand for the goods transferred in-kind, but in-kind transfers lead to a greater increase in supply in recipient markets. Using experimental data, these authors find precise indications that the prices of the goods in question decrease by about 4% in communities that receive in-kind transfers (relative to a control group). Other prices do not show significant changes. Cash transfers do not appear to have a significant effect on prices.

The remainder of the paper is organized as follows. Section 2 presents a simple model used to guide the empirical analysis. Section 3 describes the setting in which the intervention took place. Section 4 discusses the experimental design and presents the data used in this study. In Section 5, we present our empirical strategy. In Section 6, we look at the randomization balance, while Section 7 presents the empirical results. Section 8 concludes.

2. A Simple Model

Theoretical models of imperfect competition make various predictions about the competitive effects of market entry. Firms with market power may exploit their position to lower quality, just as they may raise prices (Tirole 1988). Competition attenuates this incentive, causing firms to increase quality and/or decrease prices.

---

al. (2014) design two laboratory experiments to analyze the causal effect of competition on step-by-step innovation. They find that increased competition leads to a significant increase in R&D investments by neck-and-neck firms but that it decreases the level of R&D investments by firms that are lagging behind (see also Aghion et al. (2005)).

In most models, the entry of new competitors reduces prices by putting more competitive pressure on incumbents. This is a prediction of most standard imperfect-competition models, such as differentiated-product Bertrand competition and a spatial-competition model, and also of many models with equilibrium price dispersion (such as Reinganum, 1979).

The effect of competition on product quality has been shown to be less clear-cut across the various models. Greater market power drives firms to exploit their position to increase prices and reduce quality. Competition attenuates these incentives; however firms are likely to compete through quality if quality improvements translate directly into greater demand. The effect of competition on quality depends on how well consumers perceive quality. Dranove and Satterthwaite (1992) explore the relationship between competition and quality by varying the precision of price and quality signals in a search model. They find that if competition has an effect on quality, this implies that consumers receive quality signals that are at least somewhat informative.

We now rely on a simple Cournot model of competition where n firms compete on prices and quality. We impose a reasonable set of assumptions in relation to the experiment that we analyze in this paper and derive results for the effect of competition on both the prices and quality of the goods supplied.

Assume that there are *n* equal firms that compete in a market of differentiated goods and that they choose the quantity produced (q_i) and quality (v_i), where quality could be the quality of the actual good or the quality of the delivery of the good. We will suppose that the residual inverse demand curve that a firm faces is separable in quantity and quality and that it depends not only on how much other firms supply to the market, but also on the difference between the quality of that firm's products and those of the rest of the market, as follows:

$$p_i = F\left(v_i - \alpha \sum_{j \neq i} v_j\right) - \beta q_i - \delta \sum_{j \neq i} q_j \qquad (1)$$

where F is a strictly increasing function and α is such that the argument in F is always positive. Note that the lower the value of δ, the lower the degree of substitution between goods. The same happens with the value of α. In the limit if both α and δ were zero, then increasing completion would not affect firm behavior.

For the sake of simplicity, we assume that there is no fixed cost and that the cost function is linear in the amount produced, but that it is increasing and convex in the level of quality supplied. This may reflect the fact that the initial increases in quality can be made by minor adjustments or improvements in inputs, while further improvements in quality are more costly. The cost function is then:

$$c_i(q_i, v_i) = c_q q_i + c_v \frac{(v_i)^2}{2}$$

Using both the inverse demand curve and the cost function, we can write the profit function of a firm *i* as:

$$\pi_i = \left[F\left(v_i - \alpha \sum_{j\neq i} v_j\right) - \beta q_i - \delta \sum_{j\neq i} q_j\right] q_i - c_q q_i - c_v \frac{(v_i)^2}{2}$$

The problem the firm faces is then:

$$\max_{q_i; v_i} \pi_i$$

The first-order conditions for this optimization problem are then:

$$\frac{\partial \pi_i}{\partial q_i} = \left[F\left(v_i - \alpha \sum_{j\neq i} v_j\right) - 2\beta q_i - \delta \sum_{j\neq i} q_j\right] - c_q = 0$$

$$\frac{\partial \pi_i}{\partial v_i} = F'\left(v_i - \alpha \sum_{j\neq i} v_j\right) q_i - c_v v_i = 0$$

If a symmetric Nash equilibrium exists such that $(p_i, q_i, v_i) = (p, q, v)$ for all companies, then the previous two equations become:

$$F(v[1-\alpha(n-1)]) - [2\beta + \delta(n-1)]q = c_q$$

$$F'(v[1-\alpha(n-1)])q = c_v v$$

We now use this model to investigate how the number of firms in the market affects the equilibrium values of the quantity offered and the quality chosen by each firm. In order to do this comparative static analysis, we differentiate the last two expressions to obtain:

$$\frac{dv}{dn} = \frac{(F'\alpha v + \delta q)F' - [2\beta + \delta(n-1)]F''\alpha v q}{(F')^2[1-\alpha(n-1)] - [2\beta + \delta(n-1)]\{F''[1-\alpha(n-1)]q - c_v\}} \qquad (2)$$

$$\frac{dq}{dn} = \frac{(\delta q)F''[1-\alpha(n-1)]q - (F'\alpha v + \delta q)c_v}{(F')^2[1-\alpha(n-1)] - [2\beta + \delta(n-1)]\{F''[1-\alpha(n-1)]q - c_v\}} \qquad (3)$$

If F is concave, which we will assume it is, then the denominator of the two expressions is always positive, which means that $\frac{dv}{dn} > 0$ and $\frac{dq}{dn} < 0$. In other words, as the number of firms in the market increases, the amount of the good sold by each firm in the symmetric equilibrium decreases, while the quality of the goods rises. To see how the equilibrium price reacts to entry, we then turn to equation (1) and replace the arguments with their equilibrium values:

$$p = F([1 - \alpha(n-1)]v) - [\beta + \delta(n-1)]q$$

If we differentiate this expression with respect to *n*, we get:

$$\frac{dp}{dn} = [1 - \alpha(n-1)]F' \frac{dv}{dn} - (F'\alpha v + \delta q) - [\beta + \delta(n-1)] \frac{dq}{dn}$$

Replacing $\frac{dv}{dn}$ and $\frac{dq}{dn}$ for their equivalents in equations 1 and 2 and rearranging, we arrive at:

$$\frac{dp}{dn} = \frac{[1 - \alpha(n-1)]\beta\delta(q)^2 F'' - (F'\alpha v + \delta q)\beta c_v}{(F')^2[1 - \alpha(n-1)] - [2\beta + \delta(n-1)]\{F''[1 - \alpha(n-1)]q - c_v\}}$$

From the above-mentioned assumptions, it is clear that $\frac{dp}{dn} < 0$. This means that the effect of increased competition in the symmetric equilibrium is a reduction of prices.

Thus, if customers value the increase in quality, firms, in a symmetric Nash equilibrium, will react to an exogenous increase in the number of firms by reducing prices and increasing the quality supplied. If, instead, customers do not value quality (*F*'= 0), firms will compete only on prices, as it is the case in the standard Cournot model.

3. Setting

Conditional cash transfer programs (CCTs) have been extensively used since the mid-1990s as one of the main tools for providing social protection to people in low- and middle-income developing countries. The Dominican Republic introduced a CCT program (*Programa Solidaridad*) in 2005. The program provides monetary transfers to families living in poverty. Eligibility is determined on the basis of a quality-of-life score that is used to classify households into different socioeconomic groups. All households identified as extremely-to-moderately poor are eligible.

This CCT program includes two components. First, a health component (*Comer es Primero*) provides households with a transfer of about US$ 19.5 per month.[3] In exchange, household members have to bring their children on a regular basis to the community health center for developmental monitoring and immunizations. In addition, they are expected to attend workshops that provide training on nutrition, family planning, self-care and hygiene. Second, the program also includes an educational component (*Incentivo a la Asistencia Escolar*) which gives households a certain amount of money depending on the composition of the family. Households with one or two eligible children (aged 6-16) receive US$ 8.4 per month, those with three children receive US$ 12.5, and those with four or more children

[3] Using a 2010 exchange rate of DR$ 35.9 to the dollar.

receive US$ 16.7 per month. Transfers are contingent on school enrollment and attendance of children between 6 and 16 years of age.[4] The typical household (three children of school age) would receive a total transfer of US$ 32, which represents 17% of the median monthly food expenditure of the target population.[5]

Households' monetary transfers are deposited into individual bank accounts. In order to ensure that the transfer is spent on food, the money cannot be withdrawn from the bank but, instead, can only be spent using a debit card[6] that works only in a network of retailers previously affiliated with the program (*Red de Abasto Social*), most of which are grocery stores. In those stores, program recipients can use that money to purchase only authorized consumer (food) products.[7]

The network of retailers plays a central role in this study. The process of affiliation with the network follows a standardized procedure.[8] First, the government executing agency[9] regularly opens calls for applications in certain districts and, via a community liaison, distributes application forms and encourages local stores to apply. This application procedure is, in principle, cost-free for retailers. However, many of these stores operate informally, and the application requires them to provide a tax identification number and to have a bank account. This increases the (perceived) probability of being audited. Second, interested retailers fill in and submit the application. Third, the application is reviewed and verified by the executing agency. Inspectors visit the stores and record information on the applicants' infrastructure and access to basic services, including a phone line, which is necessary in order for the debit card or magnetic stripe reader to operate and which is potentially costly for the stores. Finally, scores are assigned to the applications and stores are either granted affiliation or not depending on their score and on the number of stores already affiliated in the district. As a way of making affiliation attractive to retailers, given that there were transaction costs of operating within the network, the authorities decided to limit the number of stores that could be incorporated based on the number of beneficiaries in each district. In many districts, this could have potentially have given local market power to retailers. However, the executing agency argued that its audits had detected signs that

[4] Students must not repeat a grade more than once and must have an 85% attendance record, as a minimum.
[5] In principle, households might receive other money transfers in their bank accounts such as a subsidy for higher education (*Incentivo a la educación superior*), a pension for the elderly living in extreme poverty (*Programa protección a la vejez en pobreza extrema*), a subsidy to buy gas (*Bongas*) and/or a subsidy to pay the electricity bill (*Bonoluz*). Some of these transfers could be used in the retailers under analysis.
[6] The solidarity debit card is used only by the head of household.
[7] The CCT executing agency determined a list of products that could not be sold to beneficiaries using the debit card (e.g., alcohol). The CCT program regulations also explicitly prohibit fictitious transactions in exchange for cash.
[8] The standard process of affiliation and the operation of the retail network are governed by a set of administrative rules detailed in "Reglamento de Funcionamiento de la Red de Abasto Social" from ADESS.
[9] Social Subsidies Administration (Administradora de Subsidios Sociales (ADESS)).

consumers were being charged higher prices, that the variety of products was limited and that the quality of service was low. [10]

The particular context in which the network of retailers operates has been a cause of concern for the government in respect of two significant issues. The market power wielded by the stores belonging to the network allows them to increase prices and to offer a more limited variety of products than stores outside the network. This implies a loss of consumer surplus and therefore a potential welfare loss. In addition, over time the CCT program has been increasing the number of beneficiaries; therefore, in some parts of the country, stores currently in the network have become unable to meet the demand of the current number of beneficiaries.[11]

In response to this situation, authorities have designed a plan for the expansion of the retail network with a view to addressing the needs of all the beneficiary clientele of each store and encouraging competition among those stores in order to increase the effectiveness of the subsidies awarded under the program.

4. Experimental Design

In this context we cooperated with the CCT executing agency and the Inter-American Development Bank (IDB) to propose a way of expanding the network as a possible means of responding to the concerns raised by the government of the Dominican Republic. As part of the assistance being provided by the IDB to the government, we proposed and designed an experimental evaluation. The actual implementation of that evaluation was the responsibility of the CCT executing agency based on guidelines provided by the IDB.

The intervention consists of an exogenous increase in the number of retailers associated with the network across districts.[12] We use this randomized variability in market entry to evaluate the effect of an increase in competition on prices and retail service quality.

4.1. Experiment and Intervention

The experimental districts were identified by the CCT executing agency with two considerations in mind. They were to have, before treatment, a relatively strong demand for consumption goods per retailer, and it should be feasible, at least a priori, to expand the

[10] See report by ADESS “Proyecto de Ampliación de la Red de Abasto Social” (pp.11-13).
[11] The authorities expected to have 50 beneficiaries per grocery store in urban areas.
[12] The National Statistics Office divides the country into provinces, municipalities, sections and neighborhoods. This classification was being used by the CCT executing agency and, to simplify the project’s implementation, the evaluation was based on that same convention. Districts are defined as a collection of one or two adjacent neighborhoods.

supply of goods. Relatively high-demand districts were defined as those currently expected to have more than 100 program recipients per retailer. To increase the possibilities of expanding the product supply by recruiting new retailers, these districts should be located in municipalities with a population larger than 15,000, with at least 30% of that population being urban, and they should have at least one non-affiliated retailer that would be a priori interested in affiliating with the network. In the end, 72 experimental districts were identified. The intervention was implemented in three stages.

First, before randomization, between December 2010 and May of 2011, the CCT executing agency recruited and collected applications of retailers that wanted to become part of the network. Each one of the 72 districts was built up starting from a targeted neighborhood which was an area in which the executing agency was particularly interested in expanding the retail network. The aim was to have at least three candidates for entry in each neighborhood. However, this was not always possible, either because there were not enough applicants or because some of the applicants were not eligible. Eligibility was assessed by the executing agency on the basis of visits to the stores and audits of them; this process was contemporaneous with randomization. In those cases in which the search for potential entrants yielded few or no feasible candidates, the executing agency expanded the search area to include nearby areas (which we will call non-targeted neighborhoods). These non-target neighborhoods were adjacent to targeted areas and were also places in which, according to administrative data, beneficiaries went to do their shopping. In that sense, districts are akin to local markets. In the end, the 72 selected districts were used to provide the framework for randomization.

Table 1 provides an overview of statistics that describe the distribution of distance between retailers within districts and the distance between districts (computed using pre-intervention data). [13] The median distance between retailers within districts was about 246 meters and the median distance between districts was about 3.4 kilometers. Within the corresponding provinces, the districts are far apart.

Each district was then assigned a random number in the set {0, 1, 2, 3} which defined the number of new entrant retailers that the executing agency committed to try to affiliate. Actual affiliation could, in principle, be different from the intended/randomized affiliation because of a shortage of eligible interested candidates that applied for entry into the network (*noncompliance*). Another source of noncompliance could be the failure of the CCT executing agency to follow the intervention protocol.[14]

[13] Even though we collected information on the location of the retailers in our sample, the National Institute of Statistics of the Dominican Republic does not have the type of information that would be needed in order for us to map these neighborhoods and districts. We have therefore computed the location of the district as the centroid of the location of the retailers in our sample for each district.
[14] We performed an independent audit of compliance by calling all the retailers in the randomization sample.

Table 2 shows that, before treatment, there were some 341 retailers operating in the network within these 72 districts. Under full compliance, the design was such that a total of 99 new retailers would enter the network, which would represent an intended increase of 29% in the number of stores. A total of 21 districts were randomized to receive no entry of new retailers (*non-intention-to-treat districts*) while 51 districts were randomized {1, 2, or 3} for retailers to enter the market (*intention-to-treat districts*).

Affiliation occurred as indicated in the protocol. When the number of eligible applicants was less than or equal to the number of randomized new entrants, all of them were affiliated with the network. In those cases in which the number of applicants was larger than the number assigned by randomization, the entrants had to be selected randomly from among the eligible stores. The actual enrollment in the network was carried out in May-June 2011 by the executing agency according to a standard procedure.

Table 3 describes randomized and actual entry. A total of 61 retailers entered these districts, thereby increasing the number of firms operating in these markets by 18%. In 38 districts (53%), randomization was satisfied (*perfect compliance*), while, in 28 (39%) districts, fewer retailers than expected, according to our randomization exercise, actually entered the market (*noncompliance*), and, in 6 districts (8%), the executing agency partnered with more retailers than had originally been provided for.

4.2. Data

Baseline retailer and household data was commissioned by the IDB and collected by the *Centro de Estudios Sociales y Demográficos*, a highly respected local firm, in April and May 2011. The endline data was collected after the intervention, which took place in December 2011. Throughout the project, we also gathered administrative information from the executing agency.

We will consider three samples: the sample of retailers (both incumbent and entrants in target and non-target neighborhoods) located in the entire randomization sample of 72 districts; the sample of all retailers and consumers located in target neighborhoods within these districts; and the sample of incumbent retailers or consumers that patronize those retailers in targeted neighborhoods. Table 4 describes the sample sizes associated with each of these three sample definitions both at baseline and at endline.

The survey of retailers included the majority of incumbent retailers in the target neighborhoods (95%) and a large share of incumbent retailers in non-target neighborhoods (65%). It also provides information on all of the entrant retailers. The survey of beneficiaries was designed on the basis of a sampling frame that included all beneficiaries in the 72 target neighborhoods. The survey did not collect information on beneficiaries located in non-target neighborhoods, however. Its sample included about 30 households per neighborhood that had been drawn randomly from the sampling frame.[15]

The retailer's questionnaire was designed to collect information on the geographic location of the store; on the owner; on their participation in the retail network of the CCT; on sales, marketing and competition; and on employees and investment.

More importantly, the survey also included questions about product prices, which is one of our main outcomes of interest. During the pilot stage, we determined that, typically, only a limited number of goods were traded in these stores by program beneficiaries. These goods included bread, rice, pasta, cooking oil, sugar, flour, powdered milk, onions, eggs, beans, cod, canned sardines, chicken, salami and chocolate. For each good, we pre-specified the unit of measurement, asked owners to tell us which goods were typically available in their stores, and then asked for the price and the brand of the cheapest available option.

Since individual prices show substantial variability, in the analysis we will focus on the average price of the products sold by the retailers. This retail price is computed as the average price of items included in the survey. We have two versions of this price index: one that was computed by weighting each product by the proportion of total household expenditure (in the 15 items) measured at baseline, and another one in which a simple average was used for the computations.

The household questionnaire was designed to be answered by the person in possession of the debit card and therefore the one who did the shopping for the household. The questionnaire included queries on the physical characteristics and composition of the households, CCT program participation, the socioeconomic characteristics of the individuals living in the households, and questions on consumer behavior. Importantly, the household survey questionnaire also included a module on expenditure in which we asked about total expenditure, brands and the quantities of the same 15 items included in the retailer questionnaire. We use this to build an alternative and independent measure of the price index. For each item, we derive the price paid by the consumer from the ratio of the total expenditure on that item and the total number of units bought. Since most households did not report expenditure for all 15 items, in order to avoid a composition effect based on possible non-random non-responses on prices, we standardize each household good price by dividing it by the average price of that good as reported by all households. We then use these inferred demeaned prices to construct a weighted and an unweighted average price index, just as we did in the case of retailers. In addition the household survey includes questions that allow us to match households to retailers. We use this information to better measure the prices in the retail stores that are in our sample.

[15] The final sample has a mean and a median size of 30 per district; the smallest district has 24 beneficiaries, and the largest 60.

Note that these average prices might be affected by product quality. In order to assess the quality of the goods sold by the stores, we use the brand information gathered in the retail survey. Recall that, for each store, we have information on the price and brand of the cheapest option available at the store. For each product, we list all the brands reported in the sample and then rank them according to their average price at baseline (with a higher rank assigned to more expensive brands). We then compute a quality index per store as the average ranking of the 15 products, by store.[16] Thus, for example, if a store carries the most expensive brands for all 15 items, we infer that its average quality is higher than a store which carries the cheapest brands for all 15 items.

We are also interested in the quality of the service provided by the stores.[17] One dimension of service quality is the range of product choice available to consumers. We ask the retailers to name the three products, among the list of 15, that they sell the most of to persons using the CCT debit card. For each of these three products, we ask about the price, the variety (e.g., olive or canola oil), the brand and the unit of measurement. For each product we first rank the brands/varieties in the sample to assess product quality in same fashion as explained above. We then divide that rank by the total number of brands available in the economy. Then we compute a quality range as the (percentile) difference between the highest- and the lowest-ranked brands. Once we have computed the quality range for each product, we calculate the average quality range by store as a simple average.[18] We also measure the range of choice using the average price range by store. For each of these three products, we take the price difference between the most expensive and the cheapest available options and then compute the average price range across the three products.

To assess service quality, we also look at two direct measures. We included a question in the retailer survey about whether or not they put on special sales, and we asked consumers to score from 1 (very bad) to 10 (excellent) their last experience shopping in a retailer affiliated with the network. A last group of retailers' outcome measures are self-reported total sales, number of clients who use the CCT debit card at their store, and a dummy variable equal to 1 if the retailer perceived increased competition.

We also use administrative information by district on the total number of beneficiaries, number of retailers operating in the CCT network, and reported sales to the executing agency. Throughout the paper, we use a set of district-, beneficiaries-, and retailer-level

[16] In the endline survey, 47% of the stores reported all 15 items, 76% reported at least 14 items, and 91% reported at least 13 items. In the cases of stores that did not report all 15 items, we left that item out and computed a simple average of the reported items or a weighted average (with weight rescaled to sum to 1).

[17] We do not focus on dimensions of service quality that would require large investments, since those would probably take longer than 6 months.

[18] In this case, we did not use all 15 products but instead focused on the 8 most popular products (rice, oil, sugar, pasta, eggs, milk, beans and salami). In all, 97% of the stores reported 3 products in that set and the other 3% reported 2 products in that set.

measures as control variables to assess the validity of the design. For a full description of all these outcome and control variables, see the Data Appendix.

5. Empirical Strategy

The advantage of random assignment is that the intention to treat is exogenous. Under random assignment and perfect compliance, there is no selection into treatment status, and therefore identification of the average treatment effects is straightforward. As we have shown in Section 4.1, we do have noncompliance, especially, but not only, in districts in which the entry of two or three stores was randomized. In order to gain statistical power, we base our analysis on a parsimonious model where we pool all the treatments into a single-treatment categorical dummy variable which captures whether the district was randomized to receive one or more new stores, Z_S.

Note that, even though we had almost 50% noncompliance in the intensive margin of entry, we have better compliance when considering the extensive margin (i.e., whether there is at least one entrant into the market). Table 3 shows that in 51 districts (70%) we had entry in places randomized to entry and we observed no-entry in places randomized to no-entry. On the other hand, 21 (30%) of the districts were randomized to entry and actually observed no entry (noncompliance). Ceteris paribus, compliance was in fact better in places where there were fewer incumbent retailers before treatment. In results not shown, we estimated a logit model in which the dependent variable is a dummy of noncompliance on the extensive margin and the regressors were the original number of stores randomized to entry, a variable that indicates the number of original stores at baseline, and a variable that measures the number of beneficiaries per store at baseline. We found that having more stores randomized to entry, having a larger number of stores at baseline, and having fewer beneficiaries per store predict noncompliance with treatment. This is consistent with the idea that rents largely dissipate fast with the number of competitors in the market (Bresnahan and Reis (1991)).

Thus, in our main specifications, we estimate the following equation:

$$Y_{is} = \alpha + \gamma Z_S + \beta X_{is} + \varepsilon_{is} \qquad (4)$$

where i could be a store or a household (depending on the outcome) located in district s. Y_{is} represents any of the outcomes under study. The parameter γ captures the intention-to-treat effect of increased levels of competition on the outcome under consideration.[19] X_{is} is a

[19] Some of the variables under study are limited dependent variables (LDVs). The problem of causal inference with LDVs is not fundamentally different from the problem of causal inference with continuous outcomes. If there are no covariates or the covariates are sparse and discrete, linear models (and associated estimation techniques such as 2SLS) are no less appropriate for LDVs than they are for other types of dependent

vector of pre-treatment store- and household-level characteristics, and ε_{is} is the error term assumed to be independent across districts but allowed to display flexible correlation within districts.

Naturally, we are interested in the actual causal effect of increased competition on prices and quality.[20] Thus, we also estimate the following equation using two-stages least squares (2SLS):

$$Y_{is} = \alpha + \gamma T_S + \beta X_{is} + \varepsilon_{is} \qquad (5)$$

where T_S is a dummy variable that captures actual observed entry into the market. We instrument T_S with Z_S.

Since randomization occurs at the district level, this is our main unit of analysis. The majority of our analysis uses data at the retail or household level clustered at the district level. These standard errors are also robust to heteroscedasticity in the error term.

6. Randomization Balance

Under randomization, the outcomes of the intention- and non-intention-to-treat groups should be equal, on average, before treatment. When treatment is randomly manipulated, we have the greatest assurance that the program participants and the control group of program-eligible individuals are, on average, alike in every important sense (including observable and unobservable characteristics), with the only significant difference being that one group has been randomized into treatment and the otherwise probabilistically identical group has not. Therefore, it is common practice to test for a statistical balance of pre-treatment observable variables in order to assess the success of randomization.

Table 5 shows the mean characteristics of districts, retailers and households in the non-intention-to-treat (column 1) and intention-to-treat (column 2) groups. Column 3 shows the p-value of the null hypothesis that both means are equal. We show the balance table before treatment for all outcomes analyzed in the next section, for all variables that are included as control variables (covariates) in the models estimated in that section, and for a few other informative characteristics. Overall we observe that the mean characteristics of these groups are well balanced. We find one statistically significant difference at conventional levels out of 30 variables tested.

variables. This is certainly the case in a randomized experiment where controls are included for the sole purpose of improving efficiency, but where their omission would not bias the estimates of the parameters of interest.

[20] We do not expect general equilibrium effects to result from this experiment given that the intervention did not manipulate the transfers to the poor. Moreover, the number of markets involved in the intervention was very small relative to the whole country.

Table 5 also presents some basic statistics that provide a better picture of the setting in which the experiment took place. The districts under analysis had about 630 consumers using the CCT debit card and an average of 6 stores already operating within the retail network at baseline. Both the demand (number of beneficiaries) and the supply (retailers in the network) had been increasing in the years prior to the experiment. These characteristics are balanced across intention-to-treat groups.

The average store in our sample is a small "mom-and-pop shop"; it has about 4 employees and monthly sales of approximately USD 9,500. Usually, the owner runs and works in the store him or herself and is often joined by other family members. These grocery stores sell only food items, including raw meat and vegetables. In results not shown, the majority of the owners claim that being part of the network of retailers of the CCT has increased their sales and that about 50 percent of their consumers are beneficiaries of the CCT. All these control variables are balanced except for the number of employees, with retailers in the intention-to-treat groups having about 0.5 employees more than the average retailer in the non-intention-to-treat group.

The last panel of Table 5 describes the consumers (households) in our sample. All characteristics are balanced between the two groups. It is interesting to note that, typically, these consumers go to only one retailer to do their shopping.

7. Results

In Table 6, we present the effect of entry on log prices. Panel A shows retail prices, while Panel B shows prices as measured using household information. In the case of retail prices, we provide estimates for three samples: the whole sample, the sample of retailers located in target neighborhoods, and the sample of incumbent retailers in those target neighborhoods. In the case of households, we provide estimates for all households in the target neighborhoods and all households that bought their goods from incumbent retailers located in target neighborhoods.[21] Column 1 shows the number of observations used in the estimation and Column 2 shows the number of clusters (districts) where those observations were located.[22] Columns 3-7 show intention-to-treat estimates in which the main independent variable is a dummy for randomized entry (i.e., 1(Randomized entry>0)). Each model in those columns includes a different set of control variables, which is specified in

[21] The reader should recall that we did not collect household information in non-targeted neighborhoods.
[22] There is some variation in the number of districts/clusters across samples. Two districts only have incumbent retailers located in non-targeted neighborhoods. Therefore the sample of incumbent retailers in targeted neighborhoods has 70 clusters. Also, there is one district in which there are no consumers that buy in an incumbent retailer, so in that sample we have 71 clusters.

the bottom panel of the table.[23] Columns 8-12 show instrumental variable results in which the dummy for observed entry (i.e., 1(Observed Entry>0)) was instrumented using the randomized entry dummy. In each model we report point estimates, clustered standard errors at the district level in parenthesis, and, for the case of IV, a Klinberg-Paap F-test to assess the strength of the first-stage regression, shown between braces.

Across all samples and models, we find sizable and statistically significant decreases in prices. Since there is noncompliance, the estimates of the average causal effects are always larger than the estimates of the intention-to-treat effects. Also, consistently, for both estimands (though more pronounced in the case of the IV), the estimates are larger in absolute value for the sample of incumbent stores in target neighborhoods. The estimands are also larger for the sample of the target neighborhoods than they are for the sample as a whole. However, the effects are not statistically different.

As expected in an experimental setup like ours, adding control variables does not change the estimates noticeably. However, in our case, it does not add precision either. The estimated effects are also similar with respect to the source of information used to construct the price indexes. Moreover, as shown in Appendix Table 1, the point estimates are similar when prices indexes constructed using simple (i.e., unweighted) averages are used as the dependent variable.[24] Hence, overall, the result of entry on prices seems robust.

Regarding the size of the effect, considering the simplest IV model in Column 8, it is estimated that entry into the market decreases prices by 6% in the case of the sample of incumbent stores in the target neighborhoods. Intention-to-treat yields smaller estimates: in the same specification in Column 3, the decrease in prices is 2.5%, with the estimates not varying much across specifications. This is also consistent with having better compliance in locations with fewer incumbent retailers.[25]

In Appendix Table 2, we look at the intention-to-treat effect in districts where one store was randomized for entry and in locations where more than one store was randomized for entry. The effects are of the same order of magnitude as the ones presented in Table 6. More importantly, they are larger in districts where the entry shock is larger (i.e., where more

[23] There is not a great deal of missing data: there is complete information in all variables used in all columns for 97% of the sample of retailers.

[24] We do not have any a priori preference for using one measure (weighted) over the other (unweighted). The point estimates are similar across models and samples using both measures. The only difference is that the results for the weighted price index are more precisely estimated.

[25] In results not shown we found that sales decreased by about 9% even though these estimates were very imprecise. Taken together with the effect on prices it implies that the decrease in sales is a composition of quantities and prices reductions.

than one store was randomized to entry), although the results are not precise enough to rule out the possibility of that the estimands are equal.[26]

Table 7 presents the results on quality. The top panel shows our weighted and unweighted indexes of quality, while the bottom panel presents measures of service quality. We see quite small, insignificant effects on product quality, ranging from -2.8% to 2.1%, with many estimates bunched near zero. We interpret this result as evidence that, after entry, there was no quality change in the products sold by the stores. This result also helps us to better interpret the results on prices as a pure price effect that holds the quality of the goods constant.

The bottom panel shows the results for service quality. There seems to have been some increase in the range of products offered to the consumers even though the estimates are not statistically significant at conventional levels. The price range did not change which means that, if there was an increase in the quality range it was not done at the expense of sales of much more expensive products. Instead, the stores seem to have introduced other brands or varieties at similar prices. Also, stores in treated areas are more likely to offer special promotions and sales, and consumers rated the quality of the service higher in stores located in treated areas.

We then use the experiment to approximate a price elasticity of entry by estimating the following model:

$$\log(p_{is}) = \alpha + \delta \log\left(\frac{n_1}{n_0}\right) + \beta X_{is} + \xi_{is} \qquad (6)$$

where $\log(p_{is})$ is the log of the price index, n_0 is the number of retailers before treatment and n_1 is the number of retailers observed in the market after treatment took place. As a result of noncompliance, the causing variable (i.e. $\log(n_1/n_0)$) is potentially endogenous. Therefore, we estimate equation (6) by 2SLS using $\log\left(n_1^{Rand}/n_0\right)$ as an instrument for $\log(n_1/n_0)$, where n_1^{Rand} is the number of retailers that would have been observed under full compliance (considering the intensive margin of randomization).

The results are presented in Table 8. Using the retailer data, we find that the price elasticity of entry is about 0.08. The results are larger for incumbent retailers, which suggests that, after entry, they adjust their prices more than the entrants do. The results are a bit smaller in absolute values and more imprecise when using household-level information to measure this elasticity. However, it is nonetheless reassuring that the result holds when an independent source of information is used.

[26] We choose to show ITT effects only, rather than IV, because the first-stage F-statistics in this case are much lower (in the range of 3), mainly because we have less compliance in locations randomized to have more than one store entering the market. The results are still negative point estimates, although they become insignificant when control variables are included.

8. Conclusion

We conducted a field-randomized experiment to evaluate the effect of increased competition on prices and quality in the context of a CCT program in the Dominican Republic that provides monetary transfers to families living in poverty which can be spent only using a debit card that is not accepted anywhere except in a network of grocery stores that are affiliated with the program. The CCT executing agency was concerned that the grocery stores in the network might be capturing rents from the transfers being made to these poor households. In this context, we proposed for an expansion of the network as a possible solution for this potential problem.

Randomization was conducted at the district level. In all, 72 districts were randomized to {0, 1, 2, 3} new entrant retailers. Actual affiliation was subject to noncompliance, which was greater in the districts that were randomized to a large number of new entrant grocery stores. In order to gain statistical power, we based our analysis on a parsimonious model where we considered only the extensive margin of entry. Thus, we studied the effect of a new entry on prices and quality. We found a significant and very robust reduction of prices as a result of the increase in competition, but we did not find robust improvements in the quality of the goods or the quality of the delivery of the goods by the grocery stores six months after the intervention. We did find, however, that consumers consistently gave a higher quality rating to stores that were facing increased competition.

We then explored the impact on prices further by imposing some degree of structure. We estimated the price-elasticity of entry at 0.08. This means that, if competition increases by 1% (measured as the percentage increase in the number of firms operating in the market), then prices drop by 0.08%.

Our paper is informative for the literature on competition and efficiency. It is the first paper to provide field experimental evidence that increased competition significantly affects prices, even when the initial number of stores, on average, was not that small. As it has long been argued by economists, competition increases consumer welfare. One possible interpretation for this result, which follows from our model in section 2, is that the poor population in developing countries cares mostly about prices when shopping for groceries and is much less concerned about the types of quality dimensions that may come into play in the short-term.

Our results are also informative for the design of social policies. They suggest that policymakers should pay attention to supply conditions even when they only affect the demand side of the market. Often, social programs subsidize consumer demand by transferring resources to households. If the supply side does not operate in a very competitive environment, part of the resources targeted for the needy population will leak into the profits of the firms that are serving them. Naturally, the government could envision other options for dealing with this potential problem, as was discussed at some point in the Dominican Republic. One obvious possibility would be to attempt to regulate the market, but it has been widely recognized that the government would have to deal with an array of informational constraints in doing so. Regulation capture is another threat that has often been highlighted in the literature as an impediment to the smooth regulation of markets. Our findings, on the other hand, indicate that introducing competition provides an effective means of avoiding rent capture by suppliers.

References

Aghion, P., N. Bloom, R. Blundell, R. Griffith and P. Howitt (2005): "Competition and Innovation: An Inverted-U Relationship", *Quarterly Journal of Economics* 120, 701-728.

Aghion, P., S. Bechtold, L. Cassar and H. Herz (2014): "The causal effects of competition on innovation: Experimental Evidence", NBER WP 19987, National Bureau of Economic Research.

Basker, E., and Noel, M. (2009): "The Evolving Food Chain: Competitive Effects of Wal-Mart's Entry into the Supermarket Industry", *Journal of Economics & Management Strategy*, *18*(4), 977-1009.

Bennett, D., and Yin, W. (2013): "The Market for High-Quality Medicine", mimeo.

Berry, S. T. (1992): "Estimation of a Model of Entry in the Airline Industry", *Econometrica: Journal of the Econometric Society*, pp. 889-917.

Blundell, R., R. Griffith, and J.V. Reenen (1999): "Market share, market value and innovation in a panel of British manufacturing firms", *Review of Economic Studies* 66, 529-554.

Bresnahan, T. and P. Reiss (1990): "Entry in monopoly markets", *Review of Economic Studies,* 57(4), pp. 531-553.

Bresnahan, T. F., and Reiss, P. C. (1991): "Entry and competition in concentrated markets", *Journal of Political Economy*, pp. 977-1009.

Bresnahan, Timothy F., and Reiss, Peter C. (1989): "Do Entry Conditions Vary across Markets?", *Brookings Papers Econ. Activity*, vol. 3 (1988), pp. 833-82.

Carlton, D. W. (1983): "The location and employment choices of new firms: an econometric model with discrete and continuous endogenous variables", *Review of Economics and Statistics*, vol. 65(3), pp. 440-449.

Caves, R. E. (1980): "Industrial organization, corporate strategy and structure", *Journal of Economic Literature*, vol. 18, pp. 64-92.

Caves, R. E. and Barton D. R. (1990): *Efficiency in US Manufacturing Industries*, Cambridge MIT Press.

Caves, R.E and associates (1992): *Industrial Efficiency in Six Nations*, Cambridge MIT Press.

Cunha, J., G. De Giorgi and S. Jayachandran (2013): “The price effects of cash versus in-kind transfers”, mimeo.

Donaldson, L. (1987): “Strategy and structural adjustment to regain fit and performance: in defense of contingency theory”, *Journal of Management Studies*, vol. 24, pp. 1-24.

Dranove, D., and Satterthwaite, M. A. (1992): “Monopolistic competition when price and quality are imperfectly observable”, *The RAND Journal of Economics*, pp. 518-534.

Geroski, P. (1990): “Innovation, technological opportunity, and market structure”, *Oxford Economic Papers*, vol. 42, pp. 586-602.

Geroski, P. (1989): "The effect of entry on profit margins in the short and long run", *Annales d'Economie et de Statistique*, pp. 333-353.

Green, A. and Mayes, D. (1991): “Technical inefficiency in manufacturing industries”, *Economic Journal*, vol. 101, pp. 523-538.

Hastings, J. S., and Washington, E. L. (2008): “The first of the month effect: consumer behavior and store responses”, *National Bureau of Economic Research*, vol. w14578.

Jia, P. (2008): "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry", *Econometrica*, 76.6, 1263-1316.

Nickell, S. J. (1996): “Competition and corporate performance”, *Journal of Political Economy*, pp. 724-746.

Reinganum, J. F. (1979): “A simple model of equilibrium price dispersion”, *The Journal of Political Economy*, pp. 851-858.

Reiss, P. C., and Spiller, P. T. (1989): “Competition and entry in small airline markets”, *Journal of Law and Economics*, vol. 2, pp. 179-202.

Stigler, G. J. (1957): “Perfect competition, historically contemplated”, *The Journal of Political Economy*, vol. 65(1), pp. 1-17.

Tirole, J. (1988): *The Theory of Industrial Organization: Jean Tirole*, MIT press.

Trapani, J. M., and Olson, C. V. (1982): “An analysis of the impact of open entry on price and the quality of service in the airline industry”, *The Review of Economics and Statistics*, vol. 64(1), pp. 67-76.

Tables

TABLE 1. WITHIN- AND BETWEEN-DISTRICT DISTANCES BEFORE TREATMENT

	Number of districts	Number of retailers in the network (pre-treatment)	Distance (in mts) between retailers within districts				Distance (in mts) between districts			
			25th percentile	Median	75th percentile	Mean	25th percentile	Median	75th percentile	Mean
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
All districts	72	341	166	246	509	586	1,182	3,416	15,705	12,246
by province										
Barahona	7	11	43	83	115	80	671	928	1,114	905
Distrito Nacional	11	99	134	170	465	710	1,038	1,862	3,567	2,280
Duarte	9	31	168	222	307	240	600	796	1,320	1,077
La Vega	5	12	45	182	1,012	413	1,329	1,924	2,450	1,917
San Cristobal	6	34	191	211	606	374	855	4,719	13,472	6,397
San Pedro de Macoris	5	23	233	248	577	353	3,682	4,353	7,449	4,979
Santiago	8	31	441	576	1,042	650	1,318	2,083	2,763	2,131
Santo Domingo	17	88	222	332	589	963	5,839	34,647	40,575	24,407
Valverde	4	12	265	294	323	294	968	38,235	38,444	25,813

Note: Column [1] displays the number of districts per province. Column [2] shows the number of retailers per province. Columns [3] to [10] show the different average distances among retailers, within and between districts, in the overall districts, and per province. The values are expresed in meters.

TABLE 2. INTERVENTION AND RESEARCH SAMPLE

Number of retailers randomized for entry	Number of districts	Number of neighborhoods in each district		Number of incumbent retailers in sample
		Targeted	Not targeted	
[1]	[2]	[3]	[4]	[5]
0	21	21	6	107
1	18	18	5	71
2	18	18	6	81
3	15	15	8	82
Total	72	72	25	341

TABLE 3. RANDOMIZED AND ACTUAL ENTRY

Randomized entry (number of retailers)	Observed entry (number of retailers)					Number of districts
	0	1	2	3	4	
0	17	2	2	0	0	21
1	3	14	1	0	0	18
2	5	8	5	0	0	18
3	5	3	4	2	1	15
Number of districts	30	27	12	2	1	72

Note: Each entry shows the number of districts by randomized/observed treatment.

TABLE 4.SAMPLE DEFINITION AND SAMPLE SIZE

	Sample definition			Sample size			
Type of retailer (location)	All	Target	Incumbent in target	Retailers at baseline	Retailers at endline	Consumers at baseline	Consumers at endline
Incumbent (target)	x	x	x	215	212	1,620	1,563
Entrant (target)	x	x		42	42	630	555
Entrant (non-target)	x			9	17	-	-
Incumbent (non-target)	x			135	129	-	-
Sample size - all				401	400	-	-
Sample size - target neighborhood				257	254	2250	2118
Sample size - incumbent in target neigh.				215	212	1620	1563

TABLE 5. DIFFERENCES IN PRETREATMENT SAMPLE MEANS

	Control: No entry	Treatment: Some entry	p-value of difference	Number of obs.
	[1]	[2]	[3]	[4]
A. District Characteristics				
Log (total beneficiaries -2010)	6.441	6.453	0.960	72
	[1.016]	[0.865]		
Change in log (total beneficiaries -2009/2010)	0.211	0.172	0.380	72
	[0.200]	[0.160]		
Log (sales -2010)	11.149	11.34	0.573	69
	[1.466]	[1.165]		
Change in log (sales -2009/2010)	1.409	1.604	0.859	67
	[3.833]	[4.088]		
Number of incumbent retailers 2010	6.714	5.745	0.577	72
	[7.590]	[6.273]		
Change in log (number of retailers 2009/2010)	0.442	0.444	0.987	72
	[0.614]	[0.647]		
B. Retailer Characteristics				
Outcomes				
Log-price index - pre-treatment (weighted)	-0.323	-0.338	0.189	400
	[0.080]	[0.082]		
Log-price index - pre-treatment (unweighted)	-0.247	-0.25	0.756	400
	[0.090]	[0.078]		
Quality index	0.593	0.591	0.838	400
	[0.097]	[0.103]		
Quality index (unweighted)	0.616	0.613	0.643	400
	[0.054]	[0.060]		
Price range	0.329	0.281	0.277	361
	[0.326]	[0.307]		
Quality range	1.541	1.265	0.316	361
	[2.253]	[1.923]		
Covariates				
Log (total employees)	1.412	1.526	0.064	401
	[0.440]	[0.482]		
Percent male	0.853	0.839	0.725	401
	[0.356]	[0.368]		
1 (if the surveyed person is the retailer's owner)	0.688	0.623	0.119	401
	[0.465]	[0.485]		
1 (if has more than complete primary education)	0.679	0.613	0.197	401
	[0.469]	[0.488]		
Other variables				
1 (retailer does special sales/promotions)	0.376	0.397	0.648	401
	[0.487]	[0.490]		
Log (sales)	9.088	9.106	0.855	388
	[0.767]	[0.857]		
C. Household Characteristics				
Outcomes				
Log demeaned price (weighted)	-3.588	-3.588	0.970	2125
	[0.092]	[0.092]		
Log demeaned price (unweighted)	-3.593	-3.59	0.726	2125
	[0.100]	[0.098]		
Service quality (rating 1-10)	8.979	8.983	0.975	2248
	[1.600]	[1.413]		
Covariates				
Percent HH head male	0.635	0.618	0.620	2250
	[0.482]	[0.486]		
Household head age	53.021	52.346	0.523	2250
	[15.642]	[15.257]		
Income	475.369	498.549	0.315	2250
	[265.401]	[262.765]		
Percent of HH head married	0.576	0.538	0.293	2250
	[0.495]	[0.499]		
Percent of HH head working	0.557	0.532	0.311	2250
	[0.497]	[0.499]		
Other variables				
Amount transferred	33.804	33.609	0.805	2243
	[9.547]	[10.421]		
Number of retailers in which they shop	1.038	1.032	0.707	2249
	[0.192]	[0.176]		

Note: Columns [1] and [2] report the mean of each variable for the neighborhoods with no (randomized) entry and with some (randomized) entry. Column [3] reports the p-value of the difference test on the overall district sample. Column [4] show the number of observations used. District-cluster standard errors are reported in square brackets.

TABLE 6. IMPACT OF COMPETITION ON PRODUCT PRICES

Dependent Variable: Log (price after treatment) - weighted		Observations (number of retailers)	Clusters (number of districts)	Intention-to-treat					Average treatment effect				
				OLS estimation: 1(Entry>0) = 1(Randomized entry>0)					IV estimation: 1 (Entry>0) = 1 (Observed entry>0), instrumented with 1 (Randomized entry>0)				
		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
A) Retailer measures	*All districts*												
	1(Entry>0)	399	72	-0.020***	-0.015**	-0.015**	-0.009	-0.011	-0.040**	-0.028*	-0.028*	-0.018	-0.024
				[0.007]	[0.007]	[0.007]	[0.008]	[0.008]	[0.018]	[0.017]	[0.017]	[0.015]	[0.017]
									{10.6}	{13.6}	{13.5}	{14.1}	{13.1}
	Target neighborhoods												
	1 (Entry>0)	254	72	-0.026***	-0.019**	-0.019**	-0.019**	-0.021***	-0.056**	-0.041**	-0.041**	-0.046**	-0.056**
				[0.009]	[0.008]	[0.008]	[0.008]	[0.007]	[0.024]	[0.019]	[0.019]	[0.021]	[0.029]
									{8.4}	{8.2}	{8.3}	{6.6}	{7.0}
	Incumbent retailers in target neigh.												
	1 (Entry>0)	212	70	-0.025***	-0.019**	-0.019**	-0.019**	-0.020**	-0.060**	-0.047**	-0.047**	-0.050**	-0.061*
				[0.009]	[0.008]	[0.008]	[0.008]	[0.008]	[0.028]	[0.023]	[0.022]	[0.025]	[0.036]
									{6.5}	{6.4}	{6.4}	{5.4}	{5.7}
B) Consumer measures	*Target neighborhoods*												
	1 (Entry>0)	2025	72	-0.024***	-0.021***	-0.023**	-0.018**	-0.018**	-0.043**	-0.037***	-0.040**	-0.035*	-0.036*
				[0.008]	[0.007]	[0.009]	[0.009]	[0.009]	[0.017]	[0.013]	[0.018]	[0.020]	[0.021]
									{27.5}	{31.8}	{29.4}	{19.2}	{18.9}
	Incumbent retailers in target neigh.												
	1 (Entry>0)	1493	71	-0.030***	-0.027***	-0.022**	-0.020**	-0.021**	-0.052**	-0.047***	-0.041**	-0.041*	-0.043*
				[0.010]	[0.008]	[0.010]	[0.010]	[0.009]	[0.020]	[0.017]	[0.020]	[0.024]	[0.024]
									{26.6}	{28.9}	{22.6}	{17.2}	{17.1}
Baseline log-price index					X	X	X	X		X	X	X	X
Baseline number of retailers					X	X	X	X		X	X	X	X
Baseline quality						X	X	X			X	X	X
District controls							X	X				X	X
Retailer controls							X	X				X	X
Household controls								X					X

Note: Each entry shows an estimate of the impact of an increase in competition on the log (price) after treatment. Panel A uses the weighted log-price in the retailers database, while panel B uses the weighted log-price in the beneficiaries database. Columns [1] and [2] reports sample sizes. Column [3] and [8] report the estimation with no controls. Columns [4] and [9] control for the baseline log(price) and the baseline number of retailers. Columns [5] and [10] add the baseline quality as a control. Columns [6] and [11] report the estimates with neighborhood controls (1 (if neighborhood is targeted) and province fixed effects), and also add firm controls to the specification (owner's gender, education, and number of employees at baseline). Finally, columns [7] and [12] include household characteristics at the district level. These household controls change from panel A to panel B. Panel A includes market averages for household head's age, household income, percentage of people married, and percentage of people working. Panel B use the same controls as panel A, but computed at the individual household level. Standard errors clustered at the district level are reported in brackets. Wald F-statistic for IV estimations are included in braces. *** p<0.01, ** p<0.05, * p<0.1

TABLE 7. IMPACT OF COMPETITION ON QUALITY

(Target districts)

Outcome	Observations (number of retailers)	Clusters (number of districts)	Intention-to-treat: OLS estimation: 1 (Entry>0) = 1 (Randomized entry>0)					Average treatment effect: IV estimation: 1 (Entry>0) =1 (Observed entry>0), instrumented with 1 (Randomized entry>0)				
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Product quality												
Quality (weighted)	254	72	-0.015	-0.012	-0.010	-0.010	-0.004	-0.033	-0.027	-0.021	-0.025	-0.012
			[0.018]	[0.018]	[0.018]	[0.014]	[0.015]	[0.038]	[0.038]	[0.037]	[0.036]	[0.041]
								{8.4}	{8.2}	{8.3}	{6.6}	{7.0}
Quality (unweighted)	254	72	-0.010	-0.008	-0.006	0.002	0.001	-0.022	-0.017	-0.013	0.006	0.003
			[0.015]	[0.014]	[0.014]	[0.010]	[0.010]	[0.030]	[0.029]	[0.029]	[0.022]	[0.027]
								{8.4}	{8.2}	{8.3}	{6.6}	{7.0}
Service quality												
Quality range	235	72	0.063	0.036	-0.001	0.140	0.061	0.140	0.080	-0.003	0.344	0.172
			[0.293]	[0.297]	[0.295]	[0.209]	[0.232]	[0.630]	[0.641]	[0.637]	[0.496]	[0.629]
								{7.6}	{7.8}	{7.9}	{6.0}	{6.0}
Price range	235	72	-0.031	-0.031	-0.035	-0.008	-0.002	-0.068	-0.070	-0.076	-0.020	-0.006
			[0.075]	[0.073]	[0.073]	[0.078]	[0.074]	[0.170]	[0.167]	[0.166]	[0.186]	[0.197]
								{7.6}	{7.8}	{7.9}	{6.0}	{6.0}
1 (retailer does special sales/promotions)	254	72	0.070	0.058	0.049	0.010	0.015	0.151	0.127	0.106	0.024	0.040
			[0.080]	[0.085]	[0.084]	[0.066]	[0.070]	[0.177]	[0.186]	[0.180]	[0.153]	[0.182]
								{8.4}	{8.2}	{8.3}	{6.6}	{7.0}
Service quality (consumer measures)	2116	72	0.213**	0.234**	0.137	0.173*	0.146	0.380**	0.414**	0.253	0.302	0.291
			[0.090]	[0.089]	[0.093]	[0.098]	[0.097]	[0.192]	[0.186]	[0.178]	[0.191]	[0.189]
								{27.4}	{31.1}	{22.5}	{28.9}	{18.3}
Baseline log-price index				X	X	X	X		X	X	X	X
Baseline number of retailers				X	X	X	X		X	X	X	X
Baseline quality					X	X	X			X	X	X
Districts controls						X	X				X	X
Retailers controls						X	X				X	X
Household controls							X					X

Note: Each entry shows an estimate of the impact of an increase in competition on different outcomes. These estimates are done only for targeted districts. Columns [1] and [2] report sample sizes. Column [3] and [8] report the estimation with no controls. Columns [4] and [9] control for the baseline log(price) and the baseline number of retailers. Columns [5] and [10] add the baseline quality as a control. Columns [6] and [11] report the estimates with neighborhood controls (1 (if neighborhood is targeted) and province fixed effects), and also add firm controls to the specification (owner's gender, education, and number of employees at baseline). Finally, columns [7] and [12] include household characteristics at the district level. These household controls change depending on the outcome. The retailers' outcomes include market averages for household head's age, household income, percentage of people married, and percentage of people working as household controls. Service quality, however, uses these same controls, but is computed at the individual household level. Standard errors clustered at the district level are reported in brackets. Wald F-statistic for IV estimations are included in braces. *** p<0.01, ** p<0.05, * p<0.1

TABLE 8. PRICE ELASTICITY OF ENTRY

Dependent variable: Log (price after treatment) - weighted-		Observations (number of retailers)	Clusters (number of districts)	IV estimates				
		[1]	[2]	[3]	[4]	[5]	[6]	[7]
A) Retailer measures	*All districts*							
	$\log\left(\frac{\#Retailer_1}{\#Retailers_0}\right)$	399	72	-0.033	-0.035	-0.034	-0.046*	-0.055*
				[0.026]	[0.024]	[0.024]	[0.025]	[0.029]
				{36.1}	{34.5}	{34.9}	{39.3}	{32.5}
	Target neighborhoods							
	$\log\left(\frac{\#Retailer_1}{\#Retailers_0}\right)$	254	72	-0.052**	-0.054**	-0.054**	-0.078***	-0.091***
				[0.026]	[0.024]	[0.024]	[0.026]	[0.031]
				{26.9}	{27.1}	{27.4}	{29.5}	{24.6}
	Incumbent retailers in target neigh.							
	$\log\left(\frac{\#Retailer_1}{\#Retailers_0}\right)$	212	70	-0.077**	-0.078**	-0.078**	-0.087**	-0.103**
				[0.038]	[0.038]	[0.038]	[0.037]	[0.044]
				{19.5}	{19.8}	{20}	{17.9}	{15.9}
B) Consumer measures	*Target neighborhoods*							
	$\log\left(\frac{\#Retailer_1}{\#Retailers_0}\right)$	2025	72	-0.029*	-0.029*	-0.029*	-0.022	-0.021
				[0.016]	[0.016]	[0.017]	[0.017]	[0.017]
				{58.4}	{56.9}	{46.4}	{74.4}	{74.3}
	Incumbentr retailers in target neigh.							
	$\log\left(\frac{\#Retailer_1}{\#Retailers_0}\right)$	1493	71	-0.037**	-0.038**	-0.027	-0.021	-0.020
				[0.018]	[0.018]	[0.020]	[0.020]	[0.019]
				{36.9}	{35.8}	{30.8}	{56.1}	{55.8}
Baseline log-price index					X	X	X	X
Baseline quality						X	X	X
Districts controls							X	X
Retailers controls							X	X
Household controls								X

Note: Each entry shows an estimate of the impact of an increase in competition on the weighted price index. Panel A uses the weighted log-price in the retailers database, while panel B uses the weighted log-price in the beneficiaries database. Columns [1] and [2] report sample sizes. Column [3] reports the estimation with no controls. Column [4] controls for the baseline log(price). Columns [5] adds the baseline quality as a control. Column [6] reports the estimates with district controls (province fixed effects and total number of consumers within the district at baseline), and adds retailer's controls to the specifications (owner's gender, education, and number of employees at baseline). Finally, column [7] includes household characteristics. Standard errors clustered at the district level are reported in brackets. Wald F-statistic for IV estimations are included in braces. *** p<0.01, ** p<0.05, * p<0.1

Appendix Tables

TABLE A1. IMPACT OF COMPETITION ON UNWEIGHTED PRODUCT PRICES

Dependent Variable: Log(Price after treatment) -unweighted-		Observations (Number of Retailers)	Clusters (Number of Districts)	Intention-to-treat: OLS estimation: 1(Entry>0) = 1(Randomized entry>0)					Average treatment effect: IV estimation: 1 (Entry>0) =1 (Observed Entry>0), instrumented with 1(Randomized entry>0)				
		[1]	[2]	[3]	[5]	[6]	[7]	[9]	[10]	[12]	[13]	[14]	[16]
A) Retailer measures	*All Districts*												
	1(Entry>0)	399	72	-0.012	-0.008	-0.008	-0.002	-0.004	-0.023	-0.016	-0.016	-0.003	-0.008
				[0.008]	[0.008]	[0.008]	[0.010]	[0.009]	[0.019]	[0.017]	[0.017]	[0.018]	[0.020]
									{10.6}	{13.6}	{13.5}	{14.1}	{13.1}
	Target Neighborhoods												
		254	72	-0.029***	-0.025**	-0.025**	-0.019**	-0.023**	-0.063*	-0.055*	-0.054*	-0.045	-0.064*
				[0.010]	[0.010]	[0.010]	[0.009]	[0.009]	[0.034]	[0.032]	[0.031]	[0.030]	[0.038]
									{8.4}	{8.2}	{8.3}	{6.6}	{7}
	Incumbent Retailers in Target Neigh.												
		212	70	-0.027***	-0.024**	-0.024**	-0.018*	-0.023**	-0.066*	-0.059	-0.058	-0.047	-0.069
				[0.010]	[0.010]	[0.010]	[0.010]	[0.011]	[0.039]	[0.037]	[0.037]	[0.034]	[0.046]
									{6.5}	{6.4}	{6.4}	{5.4}	{5.7}
B) Consumer Measures	*Target Neighborhoods*												
	1(Entry>0)	2025	72	-0.024***	-0.023***	-0.020**	-0.010	-0.011	-0.042***	-0.040***	-0.035**	-0.020	-0.021
				[0.008]	[0.007]	[0.009]	[0.007]	[0.007]	[0.016]	[0.014]	[0.017]	[0.016]	[0.016]
									{27.5}	{32.7}	{29.6}	{19.4}	{19.2}
	Incumbent Retailers in Target Neigh.												
	1(Entry>0)	1493	71	-0.028***	-0.027***	-0.017*	-0.012	-0.013*	-0.049**	-0.047***	-0.032*	-0.026	-0.027
				[0.009]	[0.008]	[0.010]	[0.008]	[0.008]	[0.020]	[0.017]	[0.019]	[0.018]	[0.018]
									{26.6}	{29.6}	{22.7}	{17.4}	{17.3}
Baseline Log-Price Index					X	X	X	X		X	X	X	X
Baseline Number of Retailers					X	X	X	X		X	X	X	X
Baseline Quality						X	X	X			X	X	X
Districts controls							X	X				X	X
Retailers controls								X					X
Household controls								X					X

Note: Each entry shows an estimate of the impact of an increase in competition on the Log (price) after treatment. Panel A uses the weighted log-price in the retailer's database, while panel B weigthed log-price in the beneficiaries database. Columns [1] and [2] reports sample sizes. Column [3] and [8] report the estimation with no controls. Columns [4] and [9] control for the baseline log(price) and the baseline number of retailers. Columns [5] and [10] add the baseline quality as control. Columns [6] and [11] report the estimates with neighborhood controls (1(if neighborhood targeted) and province fixed effects), and also add firm controls to the specifications (owner's gender, education, and number of employees at baseline). Finally, columns [7] and [12] include household characteristics at the district level. These household controls change from panel A to panel B. Panel A includes market averages for household head's age, household income, percentage of people married, and percentage of people working. Panel B use the same controls as panel A, but computed at the individual household level. Standard errors clustered at the district level are reported in brackets. Wald F-statistic for IV estimations are included in braces. *** p<0.01, ** p<0.05, * p<0.1

TABLE A2. IMPACT OF ENTRY ON PRICES

Dependent Variable: Log(Price after treatment) -weighted-	(Number of Retailers)	(Number of Districts)	Intention-to-treat OLS Estimation				
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
All Neighborhoods							
1(Randomized entry=1)	399	72	-0.020**	-0.011	-0.011	0.003	0.001
			[0.009]	[0.009]	[0.009]	[0.010]	[0.009]
1(Randomized entry=2 or 3)			-0.020**	-0.016**	-0.016**	-0.014*	-0.015*
			[0.008]	[0.008]	[0.008]	[0.008]	[0.008]
Target Neighborhoods							
1(Randomized entry=1)	254	72	-0.023**	-0.014	-0.014	-0.010	-0.011
			[0.011]	[0.010]	[0.010]	[0.010]	[0.010]
1(Randomized entry=2 or 3)			-0.027**	-0.021**	-0.021**	-0.024***	-0.025***
			[0.010]	[0.009]	[0.009]	[0.009]	[0.008]
Incumbent Retailers in Target Neigh.							
1(Randomized entry=1)	212	70	-0.028***	-0.019*	-0.019*	-0.014	-0.015
			[0.011]	[0.010]	[0.010]	[0.010]	[0.010]
1(Randomized entry=2 or 3)			-0.023**	-0.019**	-0.019**	-0.022**	-0.023**
			[0.011]	[0.009]	[0.009]	[0.009]	[0.009]
Baseline Log-Price Index				X	X	X	X
Baseline Number of Retailers				X	X	X	X
Baseline Quality					X	X	X
Districts controls						X	X
Retailers controls						X	X
Household controls							X

Note: All entries report the estimation of a model in which the dependent variable is the log(price) and the independent variables are dummies indicating the level of treatment (D=1,2,3) and controls. In the cases when the observed entry was 3 or 4 we established a single dummy that takes the value of 1 if the entrance was equal to 3 or 4. Columns [1] and [2] report sample sizes. Column [3] reports the estimation with no controls. Columns [4] controls for the baseline log(price) and the baseline number of retailers. Column [5] adds the baseline quality as control. Columns [6] reports the estimates with neighborhood controls (1(if neighborhood targeted) and province fixed effects), and also adds firm controls to the specifications (owner's gender, education, and number of employees at baseline). Finally, columns [7] includes household characteristics at the district level. Standard errors clustered at the district level are reported in brackets. *** p<0.01, ** p<0.05, * p<0.1