DISCUSSION PAPER SERIES

No. 10456

ASYMMETRIC PRICE EFFECTS OF COMPETITION

Saul Lach and José-Luis Moraga-González

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

ISSN 0265-8003

ASYMMETRIC PRICE EFFECTS OF COMPETITION

Saul Lach and José-Luis Moraga-González

Discussion Paper No. 10456
March 2015
Submitted 21 February 2015

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801
www.cepr.org

ASYMMETRIC PRICE EFFECTS OF COMPETITION[†]

Abstract

In markets where price dispersion is prevalent the relevant question is not what happens to the price when the number of firms changes but, instead, what happens to the whole distribution of equilibrium prices. Using data from the gasoline market in the Netherlands, we find, first, that markets with a given number of competitors have price distributions that first-order stochastically dominate the corresponding price distributions in markets with one more firm. Second, the competitive response varies along the price distribution and is stronger at prices in the medium to upper part of the distribution. Finally, consumer gains from competition depend on how well informed they are and turn out to be larger for relatively attentive consumers. To account for these empirical results, we propose a generalisation of Varian's (1980) well-known model of sales that allows for richer heterogeneity in consumer price information.

JEL Classification: D43, D83 and L13
Keywords: distribution of price information, number of competitors and price dispersion

Saul Lach Saul.Lach@huji.ac.il
The Hebrew University and CEPR

José-Luis Moraga-González j.l.moragagonzalez@vu.nl
VU University Amsterdam

[†] We thank Marco Haan, Pim Heijnen and Adriaan Soetevent for providing us with the data and for comments. We also thank Guillermo Caruana, Riemer Faber, Chaim Fersthman, David Genesove, Alex Gerhskov, Dietmar Harhoff, Corine Hoeben, Maarten Janssen, Gerard Llobet, Jan de Loecker, Benny Moldovanu, Marielle Non, Carlos Noton, Vaiva Petrikaite, Zsolt Sandor, Dirk Stelder, Matthijs Wildenbeest, Mariano Tappata and Shlomo Yitzhaki for their remarks. The paper has also benefited from seminar participants at the ASSET Meetings 2010, Ben Gurion University, CEMFI, 10th CEPR Conference on Applied Industrial Organization, CERGE-EI, Instituto Di Tella, 24th Annual Congress of the EEA, 37th Annual Congress of the EARIE, The Hebrew University, London School of Economics, Tel-Aviv University, University of Mannheim, University of Zurich, 3rd Workshop on Search and Switching Costs and The World Bank. Boaz Abramson, Yaron Aronshtan, Danny Bahar, Eli Berglas and Sarit Weisburd provided excellent research assistance. Lach thanks the Wolfson Family Charitable Trust for its financial support.

1 Introduction

Economists have long studied the relationship between the number of firms and prices. Standard Cournot and Bertrand oligopoly models assume consumers are perfectly informed about all prices in the market and predict that an increase in the number of competitors lowers the equilibrium price. These standard models formalize the widely accepted view in economics that more competition, as measured by an increase in the number of competitors, lowers prices and benefits all consumers.

Alternative, and more realistic, models depart from the assumption that all consumers have the same information and, in the absence of the possibility to price discriminate, describe equilibria characterized by price distributions.[1] In such markets where price dispersion is prevalent the relevant question is no longer what happens to *the* price when the number of firms changes but, instead, what happens to *the whole distribution of equilibrium prices*.

In this paper we study how the distribution of prices changes with the number of competitors in a market where price dispersion is prevalent, namely, the gasoline market.

Price dispersion models tell us that, in general, an increase in competition is expected to affect different percentiles of the price distribution in a non-uniform way. This occurs because when consumers are differentially informed, an increase in the number of firms does not increase competitive pressure for all consumer groups symmetrically. In fact, as competition increases, relatively well informed consumers become disproportionately harder to attract than relatively less well informed consumers. Since firms cannot price discriminate, they are led to revise their pricing strategy in non-trivial ways. In some environments, as in the famous shopper/non-shopper formulation of Varian (1980), this asymmetric impact of competition manifests itself in a remarkable way: the "upper" portion of the price distribution shifts right while the "lower" portion shifts left, and this occurs in such a manner that the mean price goes up.[2] In more general settings, as we show later in an extension of Varian's (1980) model where we allow for arbitrary heterogeneity in consumer price information, the effect of more competition on the entire price distribution might be to lower all its

[1]The information consumers have about prices varies across them for exogenous and/or endogenous reasons. *Exogenous* reasons include the involuntary exposure to price information when consumers, for example, walk around, talk to friends and neighbours, and/or interact in social networks. *Endogenous* reasons include active consumer search as well as advertising and promotional activities. For a recent survey of models that rationalize price dispersion, see Baye et al. (2006).

[2]Varian himself did not prove this result. For a proof see Janssen and Moraga-González (2004) and Morgan et al. (2006).

percentiles.

When market equilibria are characterised by price distributions, the effects of increased competition are, therefore, potentially complex. Moreover, distinct consumers, depending on whether they are more or less alert to price information, will be affected differentially. These observations lead us to ask: How do the different percentiles of the price distribution change as competition intensifies? Do all consumers gain from competition? Which consumers gain more, those well informed about prices or those poorly informed?

To address these questions in a real-world setting, we study the gasoline market in The Netherlands. Specifically, we examine how the deciles of the price distribution vary with the extent of competition as measured by the number of firms operating in a market. We use station-level daily prices of Euro 95 gasoline obtained from 3,195 gas stations for the 8-month period March 3, 2006–October 31, 2006. In order to measure the number of competitors of a gas station i, we define the (local) market for a gas station i as the geographic area in a circle with radius r centered at the location of gas station i. We first show that all the estimated effects of an increase in the number of competitors on the deciles of the price distribution are negative and statistically significant. This implies that markets with a given number of gas stations have price distributions that first-order stochastically dominate the corresponding price distributions in markets with one more firm. Second, the estimated effects vary significantly with the decile. The estimated coefficients for the medium to upper deciles are larger (i.e., more negative) than those for the lower deciles. That is, the competitive response varies along the price distribution and is stronger at prices in the upper part of the distribution. Third, the estimated impact of an additional gas station becomes weaker as the market area increases and the typical gas station faces more competitors.

These empirical findings suggest that all consumers do benefit from an increase in the number of firms. However, because not all the deciles respond in the same way to an increase in competition, it is possible that consumer gains are asymmetrically distributed across the consumer population. In fact, we show that the largest gains from increased competitiveness accrue to consumers who are better informed about prices. Arguably, if lower income consumers have more incentives to pay attention to the posted prices while they drive around or commute, our results suggest that competition is not only efficiency enhancing but also serves to redistribute rents towards the lower income segments.

Empirical research of markets with price dispersion has usually proceeded by estimating the impact of competition on the mean and variance of prices.[3] While this is certainly informative, focusing only on these two statistics is much too narrow to uncover the clear patterns we observe in our data, namely, that more competition causes the entire distribution to shift to the left and, moreover, that price dispersion falls because the top deciles of the price distribution fall more than the bottom deciles do. Our broader approach also provides the interesting welfare implication that more competition benefits consumers unequally. In fact, consumers who are more aware of the posted prices experience greater gains from increased competition than consumers who are typically less attentive.

Varian's (1980) and Rosenthal's (1980) models fail to account for the empirical pattern we have found in the data that price distributions shift to the left as the number of competitors increases; in fact, their models predict, on the contrary, that an increase in the number of competitors raises the average price. This distinct prediction has sometimes been used to discriminate among possible explanations of observed price dispersion and Varian-type models have been dismissed as a plausible explanation on this ground (e.g., Barron et al., 2004; Hosken et al., 2008; Haynes and Thompson, 2008). One way to accommodate our empirical results is to allow for endogenous consumer search, see e.g. Dana (1994) and Janssen and Moraga-González (2004). However, we do not believe that consumer search is a prominent feature of gasoline markets because consumers hardly deviate from their commuting path in order to shop for lower prices. Instead, following Armstrong, Vickers and Zhou (2009), we propose to modify Varian's (1980) model by allowing for more general forms of consumer price information heterogeneity in order to capture consumer variation in driving and commuting patterns.[4] We study such a generalised model where we postulate that if a market has, say, 4 firms then the share of consumers who know exactly 1, 2, 3 and 4 prices is non-negative.[5] Under the weak assumption that an increase in the number of competitors lowers the share of consumers who observe one price only and increases the average number of prices observed in the market, we find plausible conditions under which the price distribution shifts to the left as competition increases.

[3]See, for example, Borenstein and Rose (1994), Barron et al. (2004), Baye et al. (2004), Lewis (2008), and Gerardi and Shapiro (2009).

[4]See also Nermuth et al. (2009), who, using a similar setting, focus on how the Internet affects the distribution of information and prices.

[5]In Varian's shopper and non-shopper formulation the assumption is that consumers observe either 1 or 4 prices but never 2 or 3.

Examples in which our theoretical model fits relatively well the observed effects in the data are when the distribution of price information in the market follows a (discrete) uniform or a (truncated) binomial distribution. In such cases, an increase in the number of firms shifts the entire distribution to the left in such a way that consumers who are more alert to the posted prices derive greater utility gains from an increase in competition.

We believe the message of this paper goes beyond the present application to the gasoline market in the Netherlands. Since imperfect price information is prevalent in many markets (telecommunications, health, gas, electricity, etc.), the price effects of competition-enhancing policies (industry deregulation, trade liberalization, etc.) might not be as straightforward as those implied by standard models. Moreover, since increased competition can potentially have unequal effects among consumers, distributional issues become a central part of the welfare assessment of these policies. This advocates the importance of taking a broader view where the interaction between competition and consumer policy is taken into consideration (Armstrong, 2008; Waterson, 2003; Wu and Perloff, 2007).

The remainder of the paper is organized as follows. Our empirical analysis of the gasoline market in the Netherlands is in Section 2. In Section 3 we propose a model, inspired by Varian's (1980) model of sales, of the distribution of prices in an oligopolistic market where consumers differ in the amount of prices they are exposed to. Conditions are derived under which prices can increase or decrease as the number of competitors goes up. Proofs are relegated to Appendix B. The paper closes with Section 4, where we offer some concluding remarks.

2 Empirical findings from the retail gasoline market

We use daily prices for Euro 95 gasoline from a large sample of gas stations in the Netherlands. The price data were obtained from Athlon Car Lease Nederland B.V., the largest private car leasing company in the Netherlands with over 129,000 cars as of the end of 2008 (www.athloncarlease.com).[6] The typical contract between Athlon and its lessees stipulates that Athlon pays for the gasoline consumed (up to a limit) as well as for car maintenance, insurance, etc. In order to do this, Athlon

[6] The data used in this paper are part of the data collected and analyzed by Soetevent *et al.* (2013). We are indebted to them for providing us with the gasoline price data and the list of gas stations operating in the Netherlands. They study whether ownership changes in highway gas stations originating from a government program of auctions and divestitures enhances competition. For further details on the data collection, see their Appendix B.

gets the lessees' gas receipts and it is from these receipts that the fuel prices are retrieved. Athlon's lessees do not get special discounts so the prices reported by Athlon are actual prices paid by drivers at the pump.

Prices were obtained from 3,195 gas stations for the 9-month period March 3, 2006–October 31, 2006. This number does not include gas stations located along highways. The sample covers 80 percent of the universe of gas stations in the Netherlands.[7] In the Netherlands, gas prices typically change on a daily basis. Because the price information arrives directly from the lessees, not all stations are sampled every day, which results in an unbalanced panel data of gas stations.

There are 306,436 station-day observations on Euro 95 prices. For illustration, Figure 1 displays the density function of prices in all gas stations in the sample. The lowest price in our sample is 102 cents while the highest price is 167. The mean and median price of Euro 95 gas is 138.9 and 140 cents, respectively, and the standard deviation is 7.15 cents.

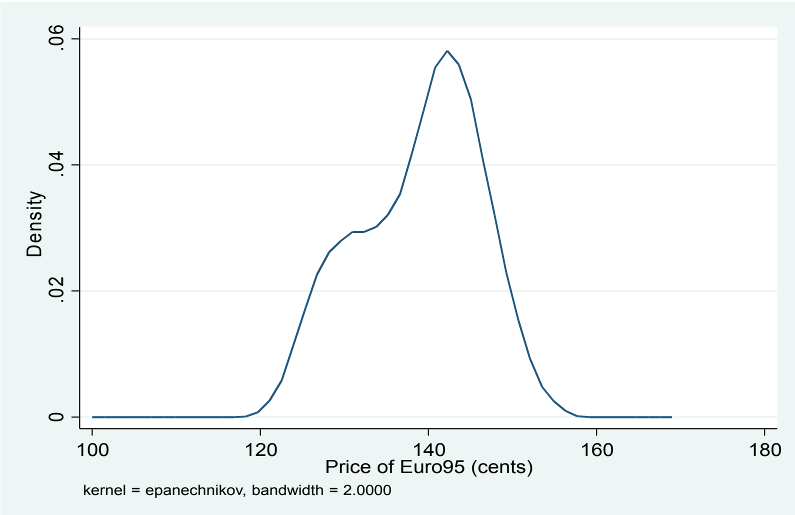

Figure 1: Kernel density estimate of Euro 95 prices (cents)

Not surprisingly, there is dispersion in gasoline prices and, although, not very large it has some economic significance. As an illustration, we computed the difference between the highest and lowest prices offered by gas stations within the same municipality and during the same day. The largest

[7]The 3,195 distinct gas stations represent, if fact, 3,260 legal entities since some gas stations changed ownerwship during the sample period. We treat these cases as a single station. The Dutch Competition Authority reports that 4,319 gas stations were registered in the Netherlands in 2005, including stations located in highways. We substract from this number 237 stations located in highways (in 2006) and therefore our data cover about 80 percent of the gas stations opearing in the Netherlands in 2006.

difference was 38 cents and occurred in the municipalities of Langedijk (on May 1) and Boxtel (on June 3). This implies that a consumer filling a 50-liter tank at the lowest-priced station instead of at the highest-priced station would have saved 19 euros.

We view our sample of prices from a gas station i as random draws from the equilibrium distribution of prices for firm i. We are able to assume this because of several reasons. First, Athlon's lessees do not pay themselves for the gas (it is part of the contract) and therefore it is reasonable to assume that they have no incentives to search for gas stations offering the lowest prices. This is important because, otherwise, our sample would have been a sample of the stations with lowest-priced gasoline.[8] Second, all gas stations in the Netherlands are self-service and therefore there is a single price for gas in each station. Finally, we believe that the extent to which various prices in a given market are set by a single firm (because of joint-ownership) and/or reflect collusive agreements is minor, implying that prices can be viewed as independent draws.[9]

We define the (local) market for a gas station i as the geographic area in a circle with radius r centered at the location of gas station i.[10] There are therefore 3,195 markets. We use the list of gas stations (and their geographic locations) to compute the number of firms N_i in market i. N_i depends on the radius r used to define the local market but we omit this from the notation. Since we have information on only 80 percent of all the stations in the Netherlands, N_i may be measured with some error. Table 1 shows the distribution of the number of gas stations N_i across markets for various radii.

As the radius increases, the number of rivals of a firm i ($N_i - 1$) increases. Note that increasing the radius from 1 to 2 km. decreases considerably the number and share of stations without competitors,

[8]On the other hand, these consumers may be choosing to buy from gas stations offering higher service values and therefore higher prices. We do not control for this type of selection but this should not be much of a problem since our sample covers about 80 percent of all gas stations in the Netherlands.

[9]Although we do not have information on the gas stations' owners, according to the Dutch Competition Authority, about 62 percent of the gas stations are owned and operated by independent dealers (NMa, 2006). The remaining stations belong to the main oil producers: BP, Esso, Shell, Texaco and Total. But even among these branded stations, most are dealer-operated. For example, Shell serves fewer than 15 percent of the gas stations and about 2/3 of the Shell-branded gas stations are operated by dealers who are free to set their own prices. This suggests that joint ownership of gas stations is not such a prevalent phenomenon as one may be led to believe from casual observation (although we have no data on joint ownership of gas stations by independent owners). An exception was the highway market, where most gas stations, 63 percent, were owned and operated by the large oil producers (NMa, 2006). However, starting in 2002, the Dutch government has forced divestitures of highway stations in order to increase competition. Stations located along highways are nevertheless not included in our sample.

[10]For an alternative definition of markets based on commuting patterns see Houde (2012). We note that we do not possess this kind of data; moreover, because we do not use gasoline stations located along highways–highways being the main commuting routes – our definition of markets is a sensible one.

Distribution of N for various market definitions				
	radius:1 km		**radius:2 km**	
Mean	1.6		3	
Median	1		3	
Std. Dev.	0.84		1.94	
N	Frequency	Percent	Frequency	Percent
1	1721	54%	746	23%
2	1000	31%	680	21%
3	368	12%	574	18%
4	86	3%	494	15%
5	15	0%	288	9%
6	2	0%	190	6%
7	2	0%	128	4%
8	1	0%	59	2%
9			24	1%
10			8	0%
11			4	0%
Total	3195	100.00%	3195	100.00%

Table 1

or with only one competitor. The median number of stations across markets varies between 1 and 3. Soetevent *et al.* (2013), using the same data, found that rivals' price effects are significant up to 2 km, and strongest at distances of 1-2 km.

The price dispersion models mentioned in the Introduction do not typically apply to monopolies (markets where $N = 1$). Correspondingly, in what follows we do not use markets in which there is just one firm. Nevertheless, note that markets with $N = 1$ become significantly less frequent as we increase the radius from 1 to 2 km. Further, price dispersion models prescribe the firms to use mixed strategies in equilibrium. Though we do not test this tenet here directly, we remark that we have done precisely this in an earlier working paper version of this paper (Moraga-González and Lach, 2008). Furthermore, Hosken *et al.* (2008), Lewis (2008), Chandra and Tapatta (2011) present evidence on price dispersion and behavior consistent with the use of mixed strategies for different samples of gasoline stations located in the US while Pennerstorfer et al. (2014) does so for the case of Austria. Hence, in what follows we focus on estimating the effect of competition on the distribution of prices.

We now proceed to estimate the effect of the number of competitors of a typical firm i on the percentiles of its distribution of prices. In order to do this, we need to control for market-specific factors that affect both prices and the number of firms in the market. For example, consumers in different markets may have distinct valuations for gasoline as well as distinct shopping behavior because of reasons as varied as differences in income, commercial activity and spatial-related factors. Operating costs may also differ across markets making some locations more expensive and consequently with higher prices and perhaps lower number of competitors. If we do not control for these factors, an omitted variable problem may arise making N_i endogenous and biasing our estimates of the effect of competition on prices.

We use fixed effects for the municipality (town) where the gas station is located to control for these market-specific factors. The gas stations are located in 440 municipalities, the majority of which are small geographic units both in terms of area and population: the median area is 62 km^2, and the median population is 22,900. It is reasonable to assume that within-municipality variation in the relevant factors affecting prices and N_i will play a lesser role the smaller the municipality. We therefore focus our analysis on the smaller municipalities. Specifically, we use the lower 2/3 portion of the municipality area distribution, i.e., municipalities which have an area less than 88 km^2.[11] We also control for brand effects since prices and number of stations may be correlated with brand. Thus, we use municipality and brand fixed effects to address the potential endogeneity of N_i.

In addition, we also pool the data over time (days). This introduces price variation over time which we want to account for in order to improve the precision of our estimates.[12] Prices vary over time mainly because the wholesale price of gasoline varies from day to day. But also shopping patterns vary over time, say, with the day of the week. We therefore use day-of-the-week dummies and a distributive lag of the wholesale gasoline price to account for these factors.

Let $q_{it}(\tau)$ be the τ^{th} percentile of firm i distribution of prices at time (day) t. We specify the following model for $q_{it}(\tau)$, which can be viewed as an approximation to the theoretical specification implied by our model in Section 3 (cf. equation (5)),

$$q_{it}(\tau) = \sum_{\ell=0}^{L} \omega_{\ell\tau} c_{t-\ell} + \delta_\tau \ln N_i + x_{it}\beta_\tau + \mu_{m(i)\tau} + e_{it\tau} \qquad \tau \in [0,1] \tag{1}$$

where $c_{t-\ell}$ is the (lagged) wholesale gasoline price at date $t-\ell, x_{it}$ is a vector of day-of-the-week and brand dummies, $\mu_{m(i)\tau}$ is a municipality fixed effect measured at the municipality m where i is

[11] We present estimates including all the municipalities as a robustness check.

[12] We are not concerned with a correlation between the (time-invariant) N_i and unobserved time-varying factors affecting prices.

located, and $e_{it\tau}$ is an error term assumed to be uncorrelated with N_i. We use the daily spot price of gasoline from the Amsterdam-Rotterdam-Antwerp (ARA) spot market to measure c_t and use 14 lags ($L = 14$) to capture the serial correlation in prices.

The logarithm approximation is a parsimonious way of capturing "diminishing returns" to an increase in N, which, as in standard oligopoly models, is a sensible assumption.

Identification of the effect of N_i on the different percentiles of the price distribution is obtained by comparing the variation in the number of competitors across gas stations within the same municipality with variation in their prices during the same day.[13] We assume that in small municipalities the variation in the number of competitors is not related to the unobserved factors in e_{it}, i.e. the component of prices driven by station-specific unit cost shocks and by differences in local market characteristics within a municipality.

We estimate the parameter δ_τ by quantile regressions methods for various values of τ, namely $\tau \in \{0.1, 0.2, ..., 0.9\}$.[14] Because we estimate each decile separately we allow the parameters in (1) to vary across deciles. Table 2 shows the estimates of δ_τ for markets defined by using radii of 1 and 2 km.[15] The estimated coefficients are proportional to the response of the τ^{th} price decile to an increase in the number of firms from N to $N+1$, $\delta_\tau \ln[1 + 1/N]$.[16] In the first two columns, we only include the 2/3 smaller municipalities in the Netherlands. The estimates reported in the last two columns include all municipalities.

A few results are worth noticing. First, all the estimated effects are negative and statistically significant. This implies that markets with fewer gas stations have price distributions that first-order stochastically dominate the corresponding price distributions with more firms. Second, the estimated effects vary significantly with τ. For the 1 km markets, the estimated coefficients clearly increase (in

[13]Most of the price variation observed in Figure 1 is within municipalities: the between-municipality variation accounts for only 6.9 percent of the total variation in price.

[14]We do not estimate more extreme quantiles because the asymptotic distribution of estimators of extreme quantiles is non-standard (see Chernozhukov, 2005).

[15]Standard errors are robust to heteroskedasticity. Ideally, one would also like to correct for any serial correlation left after accounting for 14 lags of the wholesale price, municipality fixed effects, and possible dependence in the unobserved error among gas stations. One possibility is to cluster the standard errors by municipality but this literature is in its infancy and there is no clear agreement about the correct procedure. Indeed, Hagemann (2014) is critical of existing alternatives and suggests a bootstrap procedure that, however, rules out cluster-level fixed effects as they would lead to incidental parameter problems, i.e., we would need to omit municipality-level fixed effects from the model which are critical for identification. Nevertheless, the heteroskedasticity-robust standard errors are so low that even if cluster-robust standard errors were 5 times higher (or even more) we would have rejected the null hypothesis of no effects for most parameters.

[16]The factor $\ln[1 + 1/N]$ starts at 0.69 when $N = 1$ and declines rapidly with N, being equal to 0.15 when $N = 6$.

Quantile regression estimates				
	Small markets		All markets	
	1 km market radius	2 km market radius	1 km market radius	2 km market radius
τ	δ	δ	δ	δ
0.1	-0.470***	-0.120***	-0.283***	-0.078***
	0.028	0.015	0.024	0.011
0.2	-0.441***	-0.239***	-0.218***	-0.092***
	0.033	0.017	0.023	0.011
0.3	-0.471***	-0.295***	-0.216***	-0.129***
	0.033	0.017	0.024	0.013
0.4	-0.525***	-0.311***	-0.229***	-0.156***
	0.039	0.019	0.027	0.014
0.5	-0.582***	-0.320***	-0.243***	-0.176***
	0.039	0.020	0.027	0.015
0.6	-0.654***	-0.277***	-0.287***	-0.164***
	0.041	0.022	0.031	0.016
0.7	-0.848***	-0.260***	-0.358***	-0.172***
	0.040	0.025	0.033	0.018
0.8	-1.144***	-0.157***	-0.438***	-0.138***
	0.041	0.024	0.034	0.016
0.9	-1.356***	-0.171***	-0.583***	-0.173***
	0.055	0.023	0.030	0.020
# stores	752	1,268	1,474	2,449
# observations	73,029	126,586	137,508	232,778

Note: All quantile regressions include current and 14 lags of the wholesale price of gasoline as well as brand , day-of-the-week and municipality dummies. Small markets are municipalities with an area less than 88 km^2 .
Small numerals are standard errors robust to heteroskedasticity.
* p<0.10; ** p<0.05; *** p<0.01

Table 2

absolute value) with the percentile. In this case, the competitive response is strongest at the upper part of the distribution. For the 2 km markets, the effect is non-monotonic, with the strongest impact at the center of the price distribution. Third, the estimated effects are smaller when the market area is larger (radius 2 compared to radius 1). This is to be expected because larger markets contain more gas stations on average and hence the impact of an additional gas station is weaker.[17]

For robustness purposes, we also estimate the model using all the municipalities in the Netherlands. The exogeneity assumption that ensures consistency of the quantile estimator of δ_τ is more likely to be violated if for example there are significant idiosyncratic station effects, beyond those captured by the municipality-level effects, that attract competitors and affect production costs. For example, a particular location in a large municipality may be more desirable than others due to

[17]Note, however, that the sample data underlying the estimates differs considerably when moving from radius 1 to radius 2 markets (cf. Table 1). Doing this, adds to the estimation sample 975 gas stations that were monopolies under the radius 1 definition, totalling 95,270 new observations. This represents an increase of 69 percent in the sample size.

unobservable factors, thereby attracting more stations. This would imply a positive correlation between N and e and the estimates in Table 2 would then be biased upwards, i.e. towards zero.[18] This is exactly what we observe in the last two columns of Table 2. All the estimated effects continue to be negative and statistically significant, but they are now of smaller magnitude.

It is instructive to compare our findings to those obtained from the standard approach which analyzes the effect of competition on the mean and dispersion of prices.[19] In order to do this, we follow the standard practice and regress prices on the same set of regressors used in (1), compute residuals and then regress the squared residuals on those regressors. Table 3 reports the coefficient of $\ln N$ from the mean and variance regressions.

Effect of competition on the mean and variance of prices

	Small markets		All markets	
	1 km market radius	2 km market radius	1 km market radius	2 km market radius
Mean				
Log (number of stations)	-0.787**	-0.327*	-0.407	-0.221*
	0.374	0.184	0.270	0.134
R-squared	0.912	0.900	0.911	0.902
Variance (residual squared)				
Log (number of stations)	-1.471	-0.356	-0.650	-0.730*
	1.079	0.630	0.808	0.455
R-squared	0.207	0.112	0.166	0.112
# stores	752	1,268	1,474	2,449
# observations	73,029	126,586	137,508	232,778

Note: All regressions include current and 14 lags of the wholesale price of gasoline as well as brand , day-of-the-week and municipality dummies. Small markets are municipalities with an area less than 88 km^2 .
Small numerals are standard errors clustered at the store level.
* p<0.10; ** p<0.05; *** p<0.01

Table 3

We observe that both the mean and variance of prices decline with competition; however, these effects, specially on the price variance, are not very significant. Moreover, from these results we cannot learn which parts of the price distribution are changing, and by how much, as competition increases. For example, a decrease in the mean and variance of prices is consistent with an increase in the lower deciles accompanied by a decrease in the higher deciles. In comparison, the quantile regressions in Table 2 provide this type of information and we learn that competition lowers all

[18]Introducing fixed gas station effects would solve this problem but we cannot implement this approach since N_i is constant over the sample period. The length of the sample period, 8 months, is too short to observe much entry and, in fact, we do not have any records of entry episodes in our data.

[19]See footnote 2 for examples of this approach.

deciles but in a non-uniform manner: the mid-to-higher prices decline more than the lower prices as competition intensifies.

We now proceed to assess the extent to which consumers benefit from an increase in competition. Our estimates reveal that the magnitude of the effect of increased competition on prices is not very large: an increase from 2 to 3 gas stations decreases the median price by 0.13 – 0.24 cents ($-0.32 \times \ln 1.5 - -0.58 \times \ln 1.5$), depending on the market definition, which is not a huge effect, even relative to the small observed variation in prices across gas stations (about 7 cents).

Recall, however, that the distribution of posted prices is not the same as the distribution of prices actually paid by consumers. The latter depends on the number of prices observed by the consumer and its response to competition depends on how different parts of the equilibrium price distribution are affected by it. For example, if competition has a much stronger effect on higher than on lower prices then the expected price paid is likely to decrease more for consumers observing a small number of prices than for consumers observing a large number of prices. Thus, understanding the way the whole distribution of prices changes with competition – and not only its mean and variance – is crucial for a proper assessment of the welfare implications of competition.

Because we lack data on how informed consumers are, we further analyze this issue by simulating the expected price paid by a consumer who typically observes s prices in the market of a firm i. Once we compute the paid price, we investigate how this price changes with N_i. We proceed as follows. For each local market i we have a panel data of prices for $N_i > 1$ gas stations over time (days). We draw s prices randomly and without replacement from local market i and store the minimum of these prices. This minimum price should be seen as the price paid by a consumer observing s prices. We repeat this 1,000 times for each market, and for each s, and compute the mean of these minimum prices. This is the estimate of the expected price paid when observing s prices and we denote it by $\hat{y}_{is}$.[20]

We are interested in how competition affects $\hat{y}_{is}$ for a given s and in comparing these effects across s. We therefore regress $\hat{y}_{is}$ on $\ln(N_i)$ controlling, again, for municipality fixed effects. We run a separate OLS regression for $s = 1, 2, ..., 4$, i.e., for consumers that observe up to 4 prices. We stop at 4 because there are few markets with N above 4 (see Table 1) and hence, as s increases, the

[20]The actual procedure is slightly more involved because we pool the price data over time in order to have more observations from which to sample prices and to do this we need to control for price fluctuations due to wholesale prices (details are in Appendix A).

number of observations declines dramatically. In addition, we should not expect large differences among the prices paid by consumers who compare 3 or more prices, since 3 prices is already a large number in this context. Results are in Table 4.

Expected price paid and the number of gas stations

Dep. Var: Expected price paid when observing s prices

Radius 1 km	Small markets				All markets			
Number of prices observed	s=1	s = 2	s = 3	s = 4	s=1	s = 2	s = 3	s = 4
Log (number of stations)	-0.459	-0.625*	-0.002	-0.104	-0.592***	-0.801***	-0.228	0.369
	0.329	0.325	0.352	0.629	0.229	0.22	0.31	0.662
Number of observations	752	752	261	64	1474	1474	474	106
R-squared	0.718	0.722	0.936	0.947	0.683	0.687	0.876	0.903
Radius 2 km								
Number of prices observed	s=1	s = 2	s = 3	s = 4	s=1	s = 2	s = 3	s = 4
Log (number of stations)	-0.566***	-0.751***	-0.870***	-0.813***	-0.416***	-0.623***	-0.609***	-0.601***
	0.157	0.152	0.176	0.21	0.111	0.11	0.13	0.152
Number of observations	1268	1268	945	654	2449	2449	1769	1195
R-squared	0.747	0.753	0.826	0.84	0.702	0.706	0.777	0.802

Note: All regressions include fixed muncipality effects.
Small numerals are standard errors clustered at the store level
* p<0.10; ** p<0.05; *** p<0.01

Table 4

The estimates are all negative. A negative coefficient means that the prices paid by consumers decrease as the number of competitors increases. For small markets, these effects are only highly significant when $r = 2$. This is probably due to the relatively small number of observations used in the regressions and the consequent large standard errors. Indeed, when we use all markets the effects are significant even when $r = 1$. In any case, the estimates suggest that the effect of competition increases with s so that better informed consumers benefit more from competition.[21]

To summarize our results, we note that the data clearly show that increased competition shifts the entire price distribution to the left, reducing the middle and/or higher percentiles more than the lower ones thereby resulting in a fall in the mean price and price dispersion. Our simulation results suggest that consumers who are attentive and observe various prices benefit more from increased competitiveness than inattentive ones.

The workhorse model of price competition when consumers are differentially informed is Varian

[21]When $s = 1$ the mean observed price paid equals the mean posted price and therefore the estimates in column $s = 1$ should be similar to the mean regression results in Table 3. The difference is probably due to the fact that the estimates in Table 4 are based on mean residual prices controlling for store fixed effects.

(1980). Varian's stylised shopper/non-shopper formulation of demand, however, fails to accommodate what we observe in the data because it predicts that an increase in the number of competitors increases the average price. In fact, this distinct prediction has sometimes been used to discriminate among possible explanations of observed price dispersion and Varian-type models have been dismissed as a plausible explanation on this ground (e.g., Barron et al., 2004; Hosken et al., 2008; Haynes and Thompson, 2008). Extensions of Varian-type models with endogenous consumer search can indeed generate the result that prices decrease with the number of firms. One such model has been studied in Janssen and Moraga-González (2004). However, the assumption that there is *active search* in gasoline markets is rather unrealistic because consumers hardly change their driving and commuting patterns to shop for lower prices. In the next section, therefore, we proceed to a generalization of Varian's (1980) model that generates the patterns observed in our data. The generalization relaxes the simple shopper/non-shopper formulation of Varian by allowing for more general forms of consumer information heterogeneity.

3 A model for the distribution of prices

The market for gasoline is a good example of a homogenous goods market where price dispersion is observed.[22] Part of this price dispersion is persistent and due to observed factors creating gas station differentiation, such as spatial location, brand name, service provided and the additional amenities made available at the pump. However, prices change quite frequently and it is not trivial to tell which gas station is the cheapest in a given market. After controlling for observed factors, a significant amount of price dispersion remains. This points towards there being an element of strategic price mixing. One rationale for such a mixing is that, because consumers differ in their driving and commuting patterns, and because some are more attentive than others, they end up with different amounts of price information. Some consumers are informed about one price only because they happen to drive through only one gas station while on their usual route. Other consumers drive through various gas stations, pay attention to the posted prices and choose to tank at the cheapest one. Other consumers may be less attentive and remain poorly informed even if they encounter several gas stations in their commuting path. Under these circumstances, even if gas stations are

[22]Recent empirical studies documenting price dispersion in various markets for homogeneous products include Lach (2002) and Wildenbeest (2011) for grocery products, Hortaçsu and Syverson (2004) for mutual funds, Barron, et al. (2004), Hosken et al. (2008), and Lewis (2008) for gasoline and Baye et al. (2004) for products sold online.

differentiated, the pricing equilibrium may be characterized by mixed strategies.[23]

In order to learn how consumer information heterogeneity affects pricing, we use Armstrong, Vickers and Zhou's (2009) model, which allows for a richer information structure than the earlier shopper/non-shopper formulation of Varian (1980) often used in the literature.[24] We also abstract from aspects related to gas station differentiation, but the model can easily be extended to accommodate vertical product differentiation along the lines of Wildenbeest (2011).

There are $N \geq 2$ retailers competing in prices to sell a homogeneous good to a large number L of heterogeneous consumers. At a given moment in time, a consumer wishes to purchase at most a single unit of the good.[25] The maximum willingness to pay for the good of a firm is given by v. Letting c denote the unit cost of a firm, define $k \equiv v - c$.

Following Armstrong, Vickers and Zhou (2009), for $0 \leq x \leq 1$, define

$$\alpha_N(x) \equiv \sum_{s=1}^{N} \mu_s(N) x^s \tag{2}$$

as the *probability generating function (PGF)* for the number of prices observed by consumers. The number $\mu_s(N)$ denotes the probability of observing the prices of $s \leq N$ distinct firms, and $\sum_{s=1}^{N} \mu_s(N) = 1$. The PGF (2) is meant to represent the distribution of price information in the market. As mentioned above, the main source of price information heterogeneity is consumer variation in driving and commuting patterns, as well as in attentiveness to posted prices. Nevertheless, this formulation is general enough to capture other sensible sources of information such as word-of-mouth communication and social networking (Galeotti, 2010). Note that the s^{th}derivative of $\alpha_N(x)$ with respect to (wrt) x, which we will denote $\alpha_N^{(s)}(x)$, evaluated at 0 is equal to $s!\mu_s(N)$. We shall assume that the probability of observing exactly one price is strictly between 0 and 1, i.e.,

[23]Wildenbeest (2011) presents a tractable model where firms are vertically differentiated and mix in prices. In his model, prices are serially correlated because of the vertical differentiation but at the same time because of the mixing firms positons in the price rankings change from time to time. Such model fits well what we observe in gasoline price data.

[24]Varian's (1980) model of sales is isomorphic to an all-pay auction where firms bid by cutting prices in order to win a prize consisting in the additional demand stemming from the fully informed customers (Baye et al., 1992; Moldovanu and Sela, 2001). Allowing for an arbitrary distribution of price information in the market, as we do in this paper, sets our model apart from the all-pay auction literature. First, we do have multiple heterogeneous prizes as in Barut and Kovenock (1998) but in our game a single player can win many, even all, prizes at a time. Second, since poorly informed consumers only see a few prices, a firm bidding for these consumers is only in competition with a subset of other rivals; in this sense our game is better seen as one where players participate in multiple simultaneous all-pay auctions with different number of rival players and heterogeneous prizes. To the best of our knowledge, this situation has not been studied so far.

[25]This assumption is inconsequential. All our results extend to the case where consumers have downward sloping demand functions. We assume inelastic demands to ease the exposition only.

$\alpha_N^{(1)}(0) \equiv \mu_1(N) \in (0,1)$.

Firms play a simultaneous-moves game. Let p_i be the price of a firm i. An individual firm i chooses its price taking the prices of the rival firms as given. There are no pure-strategy equilibria. To see this, consider the position of a firm i and suppose all its rivals were charging a price $\widetilde{p}$, with $c \leq \widetilde{p} \leq v$. Two forces affect price-setting of such firm i. First, there is a desire to *steal business* from its competitors, which pushes this firm to offer better deals than the rivals. This desire arises because the chance consumers see various price offers is strictly positive. Second, the possibility of *extracting surplus* from consumers who do not compare prices prompts firm i to offer higher prices than its rivals. This desire arises because there is a chance that consumers have no other option than buying at firm i. It is easy to see that either of these deviations destabilizes the proposed equilibrium price $\widetilde{p}$. Therefore a single price level cannot accommodate these two incentives.[26]

Denote the mixed strategy of a firm i by a distribution of prices F_i. We shall only study symmetric equilibria, i.e., equilibria where $F_i = F$ for all $i = 1, 2, \ldots, N$.[27] To calculate the expected profit obtained by a firm i offering the good at a price $p_i \in [c, v]$ when its rivals choose a price randomly chosen from the cumulative distribution function F, we consider the chance that firm i sells to a consumer at random. A consumer will buy from firm i if he observes the offer of firm i, which occurs with probability $s\mu_s(N)/N$, and the offer of firm i is more attractive than any other offer he receives, which happens with probability $(1 - F(p_i))^{s-1}$. The expected demand of firm i at price p_i is therefore $L \sum_{s=1}^{N} \frac{s\mu_s(N)}{N}(1 - F(p_i))^{s-1}$ which, using (2), is equal to $\frac{L}{N}\alpha_N^{(1)}(1 - F(p_i))$. The expected profit to firm i is

$$\Pi_i(p_i; F) = \frac{L}{N}(p_i - c) \cdot \alpha_N^{(1)}(1 - F(p_i)) \tag{3}$$

In a mixed strategy equilibrium, a firm i must be indifferent between offering any price in the support of F_i and offering the upper bound $\overline{p}_i$. Therefore, any price p_i in the support of F_i must satisfy $\Pi_i(p_i; F_i) = \Pi_i(\overline{p}_i; F_i)$. In symmetric equilibrium, $F_i = F$, $\overline{p}_i = \overline{p}$ and $\Pi_i = \Pi$. As a result, since $\Pi(\overline{p}; F)$ is monotonically increasing in $\overline{p}$, it must be the case that $\overline{p} = v$ and $\Pi(\overline{p}; F) = \frac{L}{N}k\alpha_N^{(1)}(0)$. Hence, F must solve

$$(p_i - c) \cdot \alpha_N^{(1)}(1 - F(p_i)) = k\alpha_N^{(1)}(0). \tag{4}$$

[26] As in Varian (1980), when $\mu_1(N) = 1$, then all firms offering $p = v$ is a pure-strategy equilibrium. When $\mu_1(N) = 0$ instead, then all firms offering $p = c$ is a pure-strategy equilibrium.

[27] Standard derivations, which can be readily adapted from, e.g., Varian (1980), show that the support of F must be a convex set and that F cannot have atoms.

for any p in $\left[c + k\alpha_N^{(1)}(0)/\alpha_N^{(1)}(1), v\right]$, the support of F.[28]

Unfortunately, equation (4) cannot be solved explicitly for F, except in special cases. Existence and uniqueness of an equilibrium price distribution can, however, be easily proven (see Burdett and Judd, 1983). Though it is in general impossible to obtain the equilibrium price distribution analytically, we can easily derive its inverse

$$q\left(\alpha_N(\tau)\right) = c + k\frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(1-\tau)} \tag{5}$$

We note that for $\tau \in [0, 1]$, equation (5) gives the τ^{th} percentile of the equilibrium price distribution of a firm.

A close look at equilibrium condition (5) serves to make an important point: what truly matters for determining the equilibrium price distribution of a firm is not N – the number of firms – but the distribution of information among consumers.[29] It is therefore changes in the distribution of information that cause more or less "competitive pressure" in the market; changes in N *per se* have no effect on prices. We summarize this simple result in:

Proposition 1 *The number of firms N affects the equilibrium price distribution only indirectly through the PGF of price information among consumers $\alpha_N(x)$.*

The model is, in fact, a model about the effect of consumers' price information on the equilibrium price distribution. Any changes in $\alpha_N(x)$ will likely affect equilibrium prices. The focus of this paper, however, is on those changes in $\alpha_N(x)$ induced by changes in N. In this regard, we make the following assumption:

Assumption 1. *An increase in the number of firms (i) (weakly) lowers the probability consumers see one price only (i.e. $\alpha_{N+1}^{(1)}(0) \leq \alpha_N^{(1)}(0)$) and (ii) raises the (weighted) average number of prices observed in the market (i.e., $\alpha_{N+1}^{(1)}(1) > \alpha_N^{(1)}(1)$).*

Assumption 1 is a rather weak and sensible assumption. In particular, we note that it is weaker than first-order stochastic dominance (FOSD).[30]

[28] Notice that the lower bound of the price distribution is always above marginal cost, which reflects the fact that firms have market power.

[29] We note that this relies on the assumption of constant returns to scale. With economies or diseconomies of scale, the decrease in quantities caused by an increase in the number of firms would have cost, and by implication, price consequences.

[30] To be sure, for our results to hold we need that either of the inequalities $\alpha_{N+1}^{(1)}(0) \leq \alpha_N^{(1)}(0)$ and $\alpha_{N+1}^{(1)}(1) \geq \alpha_N^{(1)}(1)$

3.1 Equilibrium price distribution and the number of firms

Our objective is to study the relationship between prices in a market and the number of firms. Typical studies focus on the first and second moments of the price distribution. Like in our empirical study, here, we take a broader approach and examine the response of all the percentiles of the price distribution to changes in the number of competitors.

To do this, we study the impact of a change in the PGF $\alpha_N(x)$, caused by a change in N, on the (inverse) price distribution (5). The impact of an increase in N on the percentile τ of the price distribution is

$$q\left(\alpha_{N+1}(\tau)\right) - q\left(\alpha_N(\tau)\right) = k\left[\frac{\alpha_{N+1}^{(1)}(0)}{\alpha_{N+1}^{(1)}(1-\tau)} - \frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(1-\tau)}\right]. \tag{6}$$

Since k is positive, expression (6) clearly shows that the way an increase in competition affects the different percentiles of the price distribution depends on how $\alpha_N(x)$ changes into $\alpha_{N+1}(x)$.

Proposition 2 *Suppose that the number of firms increases from N to $N+1$ and that Assumption 1 holds. Then:*

(I) There exists $\widehat{\tau} \in (0,1]$ such that all the percentiles of the price distribution below $\widehat{\tau}$ decrease.

(II) All the percentiles of the price distribution will decrease if and only if

$$\frac{\alpha_{N+1}^{(1)}(0)}{\alpha_{N+1}^{(1)}(1-\tau)} - \frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(1-\tau)} < 0, \text{ for all } \tau. \tag{7}$$

(III) If

$$\frac{\alpha_{N+1}^{(2)}(0)}{\alpha_{N+1}^{(1)}(0)} - \frac{\alpha_N^{(2)}(0)}{\alpha_N^{(1)}(0)} < 0, \tag{8}$$

then there exists $\widetilde{\tau} \in [0,1)$ such that all the percentiles $\tau \geq \widetilde{\tau}$ of the price distribution will increase.[31]

This result says that when an increase in the number of competitors does not increase the chance of consumers being informed about only one price and does raise the number of prices they know on average then more competitors in a market *always* results in a fall in the lower percentiles of the price distribution.

---

is strict. We have chosen to work with the second inequality being strict but our proofs can easily be adapted to the similar case where the first inequality is strict instead.

[31] In special cases, it may happen that $\alpha_{N+1}^{(2)}(0) = 0$. In those cases, we can invoke higher order derivatives. In particular, the condition would involve the lowest $s \geq 2$ for which $\alpha_{N+1}^{(s)}(0) > 0$ (see the Appendix B).

Condition (7) is a necessary and sufficient condition for the equilibrium price distribution with N firms to dominate in a FOSD sense the distribution with $N+1$ firms. In such a case, *all* percentiles of the price distribution fall as we move from an N- to an $N+1$-firm market. This situation accords with the usual intuition that markets with more firms have lower prices and we remark that (7) is violated in Varian's (1980) model.[32] For an easier-to-interpret sufficient condition, we can state,

Corollary 1 (of Proposition 2) *A sufficient condition for all the percentiles of the price distribution to decrease is that the PGF $\alpha_N(x)$ satisfies:*

$$\frac{\alpha_{N+1}^{(s)}(0)}{\alpha_{N+1}^{(1)}(0)} - \frac{\alpha_N^{(s)}(0)}{\alpha_N^{(1)}(0)} \geq 0 \text{ for all } s = 1, 2, ..., N$$

We note that this condition is weaker than the monotone likelihood ratio property (MLRP).[33] In other words, if the probability distribution of price information satisfies the MLRP then increased competition implies a fall in all the percentiles of the equilibrium price distribution.

The last part of Proposition 2 says that the upper percentiles of the price distribution will increase if condition (8) holds. Note that cutting prices to capture well informed consumers results in lower expected profits for the firms. As a result, firms try to compensate by adjusting the frequency with which they charge higher prices, thereby generating higher profits from the consumers who are less well informed about prices. As we move up in the price distribution, the prices are less and less successful at capturing well informed consumers. In effect, at the top of the price distribution, firms only care about consumers observing one or two prices because the chance of selling to other (better informed) consumers is negligible. Given this, if the probability that consumers observe one price relative to the probability that consumers observe two prices increases when we move from an N- to an $N+1$-firm market, then firms prefer to raise the frequency of the higher prices and so the upper percentiles of the price distribution increase.

Proposition 2 has stated conditions under which the percentiles of the price distribution increase or decrease when we move from a market with N retailers to a market with $N+1$ retailers; however, the proposition is silent with respect to whether some percentiles increase (or decrease) more than others. As we have seen in our empirical application, this is a relevant issue because an increase

[32] In fact, in Varian's model $\mu_1(N) = \alpha_N^{(1)}(0) = \alpha_{N+1}^{(1)}(0) = \mu_1(N+1)$ and $\mu_N(N) = 1 - \alpha_N^{(1)}(0) = 1 - \alpha_{N+1}^{(1)}(0) = \mu_{N+1}(N+1)$. In this case, condition (7) requires $N - (N+1)(1-\tau) < 0$, which can never be satisfied for all $\tau \in [0, 1]$.

[33] The MLRP requires that the ratio $\alpha_N^{(s)}(0)/\alpha_{N+1}^{(s)}(0)$ decreases in s (see Milgrom, 1981).

in the number of competitors may be felt more in some percentiles than in others and this opens up the possibility that consumers' gains from an increase in competition be asymmetric in sign and magnitude. To investigate this issue further, we analyze in detail some examples.

Example 1 (The truncated binomial distribution) [34] *The truncated binomial distribution (TBD) has PGF*

$$\alpha_N(x) = \frac{[1-p(1-x)]^N - (1-p)^N}{1-(1-p)^N}.$$

where $p \in [0,1]$ is the success probability of a Bernoulli experiment. The experiment consists of observing (or not) a price and the binomial distribution gives the probability of observing s prices out of N independent trials. Note that Assumption 1 holds for the TBD. In fact,

$$\alpha_N^{(1)}(0) = \frac{Np(1-p)^{N-1}}{1-(1-p)^N}$$

Taking the derivative of $\alpha_N^{(1)}(0)$ wrt N gives

$$\frac{p(1-p)^{N-1}\left[1-(1-p)^N + N\ln(1-p)\right]}{\left[1-(1-p)^N\right]^2}. \tag{9}$$

The sign of (9) depends on $1-(1-p)^N + N\ln(1-p)$, which decreases in p. Setting $p=0$ in this expression gives 0, which implies that (9) is always negative. As a result $\alpha_N^{(1)}(0)$ decreases in N (first part of Assumption 1). Moreover, the mean of the TBD is $Np/\left[1-(1-p)^N\right]$. Taking the derivative of the mean wrt N gives

$$\frac{p - p(1-p)^N\left[1-N\ln(1-p)\right]}{\left[1-(1-p)^N\right]^2}. \tag{10}$$

The sign of (10) depends on $1-(1-p)^N\left[1-N\ln(1-p)\right]$, which increases in N. Setting $N=2$ in this expression gives $1-(1-p)^2\left[1-2\ln(1-p)\right] > 0$ for all p. Hence (10) is always positive and therefore the mean increases in N (second part of Assumption 1).

We now consider condition (7). We have that

$$\frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(1-\tau)} = \frac{\frac{Np(1-p)^{N-1}}{1-(1-p)^N}}{\frac{Np(1-p\tau)^{N-1}}{1-(1-p)^N}} = \left(\frac{1-p}{1-p\tau}\right)^{N-1},$$

which clearly decreases in N. As a result, when the distribution of price information in the market follows the TBD, an increase in the number of competitors leads to lower (in a FOSD sense) prices.

[34] We refer here to the zero-truncated distribution, since our random variable –the number of prices consumers observe– has support $\{1,2,...,N\}$.

In order to understand whether lower percentiles fall more in N than higher percentiles, we take the derivative of (6)

$$\frac{\alpha_{N+1}^{(1)}(0)}{\alpha_{N+1}^{(1)}(1-\tau)} - \frac{\alpha_{N}^{(1)}(0)}{\alpha_{N}^{(1)}(1-\tau)} = \left(\frac{1-p}{1-p\tau}\right)^{N} - \left(\frac{1-p}{1-p\tau}\right)^{N-1}$$

wrt τ, which gives

$$\left(\frac{1-p}{1-p\tau}\right)^{N-2} \frac{p(1-p)}{(1-p\tau)^3} \left[1 - Np(1-\tau) - p\tau\right]. \tag{11}$$

Inspection of (11) reveals that its sign is always positive provided that $\tau > (Np-1)/\left[p(N-1)\right]$. From this we conclude that when $p < 1/N$, (11) is positive for all τ, hence the lower percentiles fall more than higher percentiles. When $p > 1/N$, then there exists a critical $\breve{\tau}$ such that the fall in the percentiles increases in $[0, \breve{\tau}]$ and decreases in $[\breve{\tau}, 1]$. Figure 1 presents two examples.

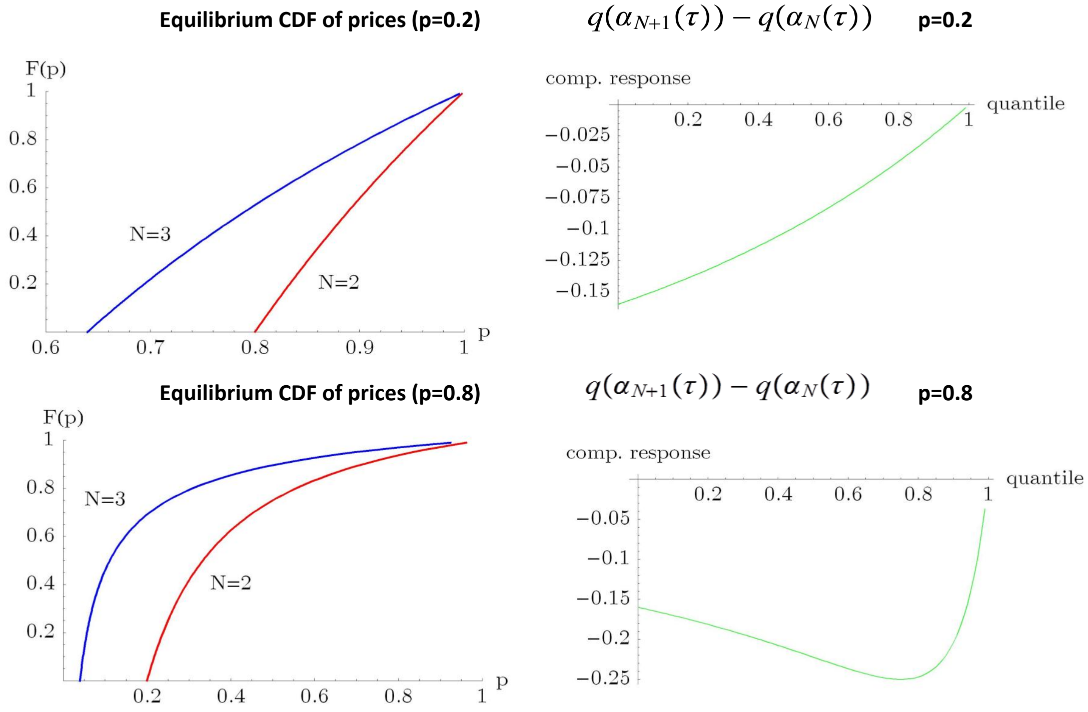

Figure 2: The truncated binomial distribution

Example 2 (The discrete uniform distribution) *Suppose consumers are equally likely to observe $1, 2, \ldots, N$ prices. The probability distribution of observing s prices is therefore a discrete uniform distribution (UD). Its PGF is*

$$\alpha_N(x) = \frac{x(1-x^N)}{N(1-x)}.$$

We first note that Assumption 1 also holds for the UD. In fact, $\alpha_N^{(1)}(0) = 1/N$, which decreases in N. Moreover, the mean of the UD is $(N+1)/2$, which increases in N.

Regarding condition (7), we have

$$\frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(1-\tau)} = \frac{\tau^2}{1-(1+N\tau)(1-\tau)^N} \tag{12}$$

Taking the derivative of (12) wrt N gives

$$\frac{(1-\tau)^N \tau^2 \left[\tau + (1+N\tau)\ln(1-\tau)\right]}{\left[1-(1+N\tau)(1-\tau)^N\right]^2} \tag{13}$$

The sign of this expression is equal to the sign of $\tau + (1+N\tau)\ln(1-\tau)$, which decreases in τ. Setting $\tau = 0$ in this expression gives 0, hence (13) is always negative. As a result (12) decreases in N and hence condition (7) holds. We conclude that when the distribution of price information in the market follows the discrete uniform distribution, an increase in the number of competitors leads to lower (in a FOSD sense) prices.

In connection with the question which percentiles decrease more, we note that the derivative of the RHS of (6) wrt τ is rather difficult in this case. However, if we set $N = 2$, such derivative gives

$$\frac{8-6\tau}{[6+\tau(3\tau-8)]^2} - \frac{2}{(3-2\tau)^2} > 0 \text{ for all } \tau.$$

Therefore we conclude that when $N = 2$, the lower percentiles fall more than the higher percentiles. The top panel in Figure 2 presents an example. We have explored numerically the cases $N \geq 3$ and found that for these cases there exists a critical $\check{\tau}$ such that the fall in the percentiles increases in $[0, \check{\tau}]$ and decreases in $[\check{\tau}, 1]$. See bottom panel of Figure 2 for an example.

Example 3 (Varian's (1980) information structure) *Varian's (1980) information structure has PGF*

$$\alpha_N(x) = \mu x + (1-\mu)x^N,$$

for some $0 < \mu < 1$. Note that $\alpha_N^{(1)}(0) = \mu$, which is constant in N, and the mean is $\mu + N(1-\mu)$, which increases in N. Therefore, Assumption 1 holds.

Regarding condition (7), we have

$$\frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(1-\tau)} = \frac{\mu}{\mu + N(1-\mu)(1-\tau)^{N-1}} \tag{14}$$

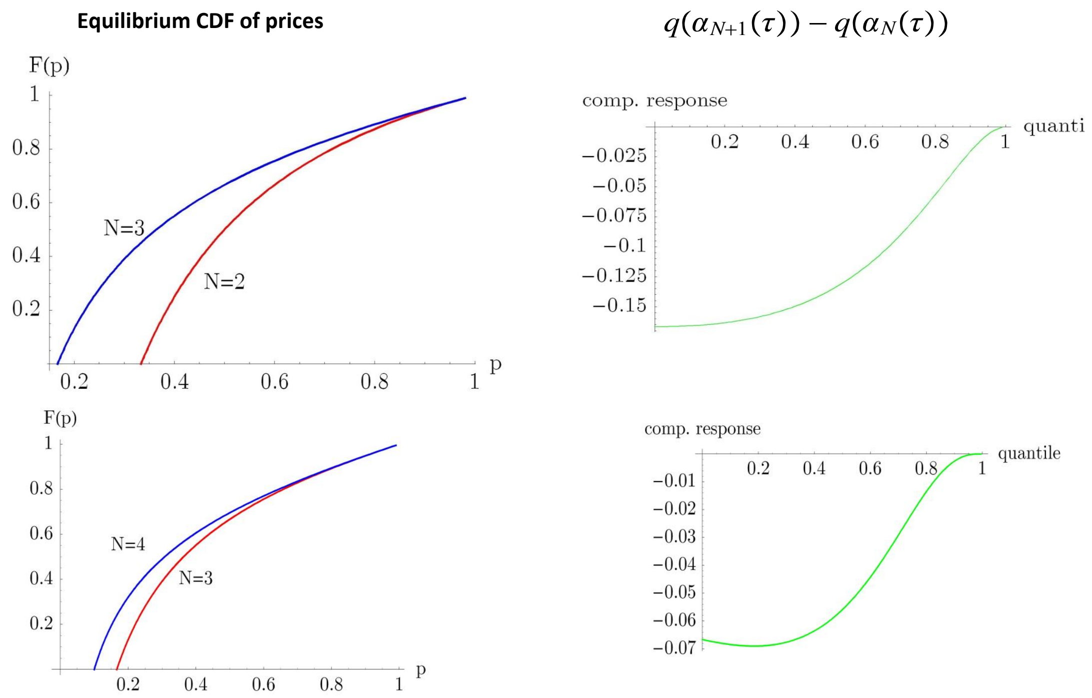

Figure 3: The discrete uniform distribution

Taking the derivative of (14) wrt N gives

$$-\frac{\mu(1-\mu)(1-\tau)^{N-1}\left[1+N\ln(1-\tau)\right]}{\left[\mu+N(1-\mu)(1-\tau)^{N-1}\right]^2}. \tag{15}$$

The sign of this expression is the opposite of the sign of $1+N\ln(1-\tau)$, which is positive for $\tau \leq \exp[-1/N]$ and negative otherwise. As a result, (15) is negative for low τ and positive for high τ. We conclude that condition (7) is violated.

Consider now condition (8). Since $\alpha_{N+1}^{(s)}(0) = 0$ for all $s = 2, ..., N$, we invoke the $(N+1)^{th}$ derivative of $\alpha_{N+1}(x)$. Then we have

$$\frac{\alpha_{N+1}^{(N+1)}(0)}{\alpha_{N+1}^{(1)}(0)} = \frac{(N+1)!(1-\mu)}{\mu} > 0,$$

which implies that condition (8) holds. We conclude that when the distribution of price information in the market follows Varian's distribution, an increase in the number of competitors leads to an increase in the high percentiles of the price distribution, and to a decrease in the low percentiles of the price distribution. The top panels in Figure 3 provide an example (for $\mu = 0.5$).

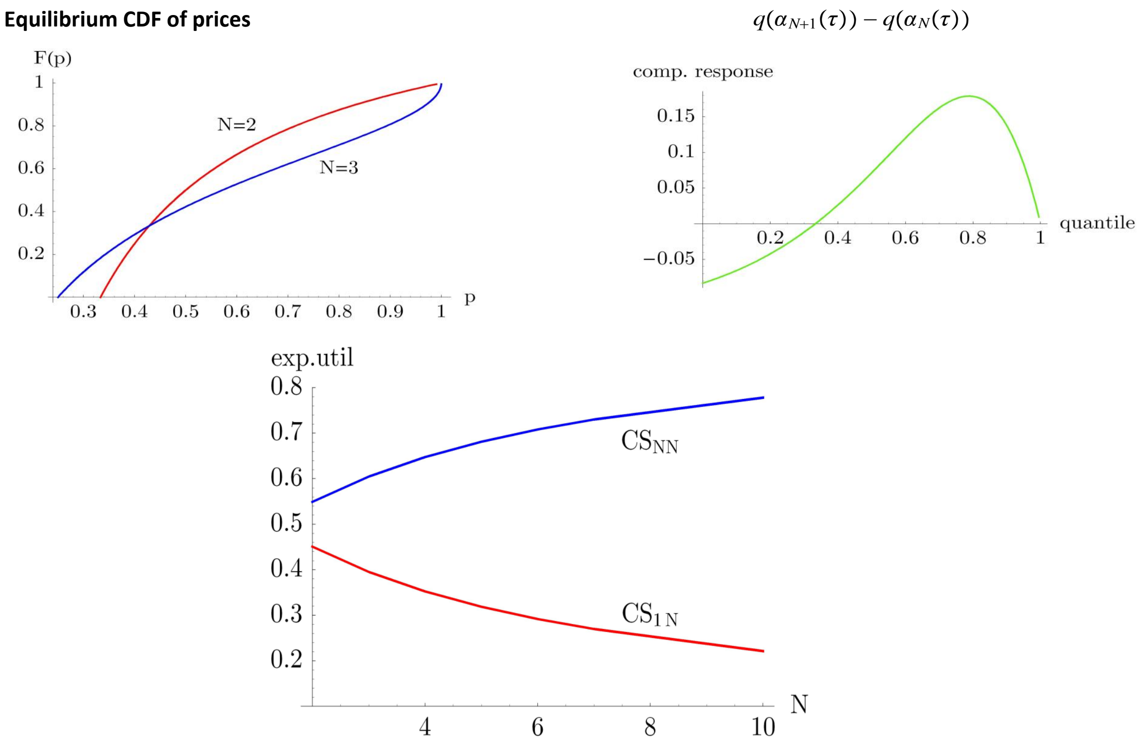

Figure 4: The shopper/non-shopper distribution ($\mu = 0.5$)

3.2 Consumer welfare and the number of firms

We have seen that the response to an increase in competition differs across the percentiles of the price distribution. In particular, we have shown (i) that some percentiles may increase while others decrease, and (ii) that if they all fall, some may decrease more than others, depending on the properties of the PGF $\alpha_N(x)$. These two results have important implications. Suppose, first, that all percentiles decrease when the number of competitors increases from N to $N+1$. Because some prices may decline more an others, it is possible that consumers observing a given number of prices derive greater benefits from increased competition than other consumers observing a different number of prices. Second, because the frequency of high prices can actually increase, some consumers may end up paying higher prices, even on average, after an increase in the number of competitors. These implications of the model are in stark contrast to standard full-information oligopoly models. In this subsection we proceed to study the welfare gains that different consumers (i.e., consumers observing different number of prices) will derive from an increase in competition.[35]

[35]Note that we ignore non-price effects of competition such as better quality of service, shorter distances to retailers, etc., so that "welfare effects" here refers exclusively to price effects.

The utility of a consumer who buys from a firm i at a price p_i is given by $v - p_i$. Denote the utility of a consumer who observes s prices and buys from the cheapest seller by $u_s = \max\{v - p_1, v - p_2, ..., v - p_s\}$ where $p_1, p_2, ..., p_s$ are i.i.d. random variables drawn from the equilibrium price distribution F. The distribution of u_s is $(1 - F(v - u))^s$. Using the price distribution from Section 2, we can derive the inverse of the distribution of u_s:

$$y_s(\alpha_N(\tau)) = k\left[1 - \frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(\tau^{1/s})}\right]. \tag{16}$$

where $\tau \in [0, 1]$. In this case, for a given τ, (16) gives the τ^{th} percentile of the distribution of the maximum utility received by a consumer who observes s prices. Because v is fixed this distribution provides the same information as the distribution of prices paid and we will also informally refer to (16) as the τ^{th} percentile of the distribution of prices paid.

Following the same steps as before, we can study how the distribution of utilities received by a consumer who observes s prices changes when we move from an N-firm to an $N+1$-firm market. We have

$$y_s(\alpha_{N+1}(\tau)) - y_s(\alpha_N(\tau)) = k\left[\frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(\tau^{1/s})} - \frac{\alpha_{N+1}^{(1)}(0)}{\alpha_{N+1}^{(1)}(\tau^{1/s})}\right] \tag{17}$$

Note the similarity between the expression in (17) and that in equation (6). In our empirical section we take advantage of this equation in order to estimate the impact of an increase in N on the utility distribution of a consumer observing s prices. Comparing (17) and (6), it is immediate to see that the change in the percentile τ of the price distribution is equivalent to the change in the percentile $(1-\tau)^s$ of the utility distribution of a consumer observing s prices.

We then have,

Corollary 2 (of Proposition 2) *Suppose that the number of firms increases from N to $N+1$ and that Assumption 1 holds. Then, for all $s = 1, 2, ..., N$:*

(I) There exists a percentile $\widehat{\tau}_s \in [0, 1)$ such that all the percentiles $\tau > \widehat{\tau}_s$ of the distribution of utilities received by a consumer observing s prices increase.

(II) All the percentiles of the distribution of utilities received by a consumer observing s prices increase if and only if condition (7) holds.

(III) If condition (8) holds, then there exists $\widetilde{\tau}_s \in (0, 1]$ such that all the percentiles below $\widetilde{\tau}_s$ of the distribution of utilities received by a consumer observing s prices decrease.

In line with Proposition 2, an increase in the number of firms results in an increase in the upper percentiles of the distribution of utilities derived by all consumers. When condition (7) holds, then all the percentiles of the utility distribution of a given consumer group increase after the number of firms goes up. Under condition (8), the distribution of the price paid by any type of consumer when there are $N+1$ firms in the market surely "crosses-over" the distribution function when there are N firms. In this situation, increased competition raises the probability that consumers who pay a high price are worse off, despite Assumption 1 being satisfied.[36]

Corollary 2 tells us about how the distribution of the prices paid by a consumer observing s prices changes with the number of firms. To study whether expected consumer welfare increases or decreases, and whether this depends on the number of prices consumers observe, we first compute expected utility conditional on observing s prices, namely

$$CS_{sN} = k\left[1 - \int_0^1 \frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(\tau^{1/s})} d\tau\right], \tag{18}$$

and then study the sign of $CS_{sN+1} - CS_{sN}$, which is equal to

$$CS_{sN+1} - CS_{sN} = k\int_0^1 \left[\frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(\tau^{1/s})} - \frac{\alpha_{N+1}^{(1)}(0)}{\alpha_{N+1}^{(1)}(\tau^{1/s})}\right] d\tau \tag{19}$$

We then have,

Proposition 3 *Suppose that the number of firms increases from N to $N+1$ and that Assumption 1 holds. Then:*

(I) If (7) holds, all consumers (i.e., for all $s = 1, 2, \ldots, N$) derive greater expected utility conditional on s.

(II) If (8) holds and $\alpha_N^{(1)}(0) = \alpha_{N+1}^{(1)}(0)$, consumers observing one price only derive lower utility.

This result states that, under condition (7), all consumers will obtain greater expected utilities given the number of prices observed. The proposition however does not inform us about whether some consumers benefit more than others. In fact, it is quite difficult to evaluate analytically how (19) depends on s. Using the examples above, however, we can gain some insights into this issue. For example, for the TBD case with $p < 1/N$, the lower percentiles fall more than the higher ones so we

[36] Anecdotal evidence tells us that many consumers report not to have felt the "supposed gains" from increased competition in liberalized markets such as airlines, gasoline, telecoms, etc. This might be related to this fact in combination with consumers remembering bad news (about prices) more readily than good news.

expect utility gains from increased competition to rise in s. By contrast, when $p > 1/N$, the impact of increased competition is felt more at intermediate percentiles of the equilibrium price distribution and as a result we expect utility gains from increased competition to fall in s.

Example 4 (The truncated binomial distribution (cont'd)) *When the price information consumers have follows the TBD, the utility gains from increased competition, equation (19), are as follows*

	Expected Utility Gains ($p = 0.2$)		
	$N = 2$	$N = 3$	$N = 4$
$s = 1$	*0.092*	*0.080*	*0.069*
$s = 2$	*0.118*	*0.101*	*0.085*
$s = 3$		*0.109*	*0.092*
$s = 4$			*0.095*

	Expected Utility Gains ($p = 0.8$)		
	$N = 2$	$N = 3$	$N = 4$
$s = 1$	*0.202*	*0.080*	*0.037*
$s = 2$	*0.198*	*0.061*	*0.021*
$s = 3$		*0.051*	*0.015*
$s = 4$			*0.012*

In this case, gains from increased competition increase in s when p is small and decrease in s when p is high.

Example 5 (The discrete uniform distribution (cont'd)) *When the price information consumers have follows the UD, the utility gains from increased competition, equation (19), are as follows*

	Expected Utility Gains		
	$N = 2$	$N = 3$	$N = 4$
$s = 1$	*0.114*	*0.045*	*0.022*
$s = 2$	*0.143*	*0.058*	*0.029*
$s = 3$		*0.064*	*0.032*
$s = 4$			*0.034*

The table shows that better informed consumers gain more from increased competition.

These two examples show two important points: (i) consumer benefits from increased competition depend on s and therefore utility gains are asymmetrically distributed across the consumer population; and (ii) utility gains need not be increasing in s, that is, poorly informed consumers may benefit more that well informed ones.

Proposition 3 also states the remarkable result that some consumers may experience a welfare loss after the number of competitors increases. In fact, when condition (8) holds, $\alpha_{N+1}^{(1)}(0) = \alpha_N^{(1)}(0)$ suffices for those consumers observing one price only to experience a welfare loss. This explains why mean prices increase in N in Varian (1980).

Example 6 (Varian's distribution (cont'd)) *When the distribution of price information among consumers follows Varian's distribution then the utility gains from increased competition are negative for the consumers who observe one price only and positive for the consumers who observe all prices in the market. The bottom panel in Figure 3 provides and example for $\mu = 0.5$.*

4 Conclusions

In markets where the amount of price information varies across consumers, prices are typically dispersed in equilibrium. When this happens, the relevant question is no longer what happens to the price when the number of firms changes but, instead, what happens to the whole distribution of equilibrium prices. Using data from the gasoline market in the Netherlands, we have found, first, that markets with a given number of competitors have price distributions that first-order stochastically dominate the corresponding price distributions in markets with one more firm. Second, our data have revealed that the competitive response varies along the price distribution and is stronger at prices in the medium to upper part of the distribution. Finally, we have performed simulations to uncover how the gains from increased competition vary as consumer price information changes. It turns out that consumers who are more alert to the posted prices by the firms consumers derive greater gains from competition than those who are less attentive and stay poorly informed.

To account for these empirical results, we have proposed a generalisation of Varian's (1980) well-known model of sales that allows for richer heterogeneity in consumer price information. The model makes clear that increased competition has an effect on prices only when it changes the amount of price information consumers have. Though the generalisation increases the complexity of the model, we have shown that it can generate the observed patterns in the data. We view this as an important observation because Varian's model distinct predictions have sometimes been used to dismiss it as a possible explanation for observed price dispersion in real-world markets.

Since price dispersion is prevalent in many markets, we believe the paper has a general message that goes beyond the present application to the gasoline market in the Netherlands. The price effects of competition-enhancing policies (e.g., industry deregulation, trade liberalization, etc.) are not as straightforward as one may be led to believe based on standard oligopoly theory. As a result, welfare implications are not obvious either. In fact, we have shown, theoretically and empirically, that increased competition can have unequal effects among consumers; at least theoretically, some consumers may even experience declines in their welfare as a result of some prices going up.

In order to identify which consumers benefit more and which benefit less from increased competition in gasoline prices we would require a mapping between shopping behavior and socio-economic characteristics of interest. If such data were available we would be able to say something about how the distribution of the benefits from increased competition varies with income. For example, consumers that observe only a few prices may be high-income consumers (whose value of time is higher) and these consumers may benefit less from competition than low-income consumers. This, however, is beyond the scope of this paper and is left for future research. Moreover, a complete welfare analysis, should recognize the effect of increased competition on other dimensions of consumer welfare such as increased variety, quality and accessibility.

Lastly, we think the empirical findings reported in the paper and that have led us to extend Varian's model are of interest on their own right and, if verified in other data sets, they should be taken into account when formulating theoretical models of pricing in oligopolistic markets.

References

[1] Armstrong, Mark (2008), "Interactions between Competition and Consumer Policy," *Competition Policy International* 4(1), 97-147.

[2] Armstrong, Mark, John Vickers and Jidong Zhou (2009), "Consumer Protection and the Incentive to Become Informed", *Journal of the European Economic Association* (Papers and Proceedings) 7(2-3), 399-410.

[3] Barut, Yasar and Dan Kovenock (1998), "The Symmetric Multiple Prize All-Pay Auction with Complete Information," *European Journal of Political Economy* 14, 627-644.

[4] Baye, Michael. R., Dan Kovenock, and Casper de Vries (1992), "It Takes Two to Tango: Equilibria in a Model of Sales," *Games and Economic Behavior* 4, 493510.

[5] Baye, Michael. R., John Morgan and Patrick Scholten (2004), "Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site", *The Journal of Industrial Economics* 52(4), 463-496.

[6] Baye, Michael. R., John Morgan and Patrick Scholten (2006), "Information, Search, and Price Dispersion, chapter 6 in *Handbook on Economics and Information Systems Volume 1,* (T. Hendershott, Ed.), Amsterdam: Elsevier.

[7] Barron, John M., Beck A. Taylor and John R. Umbeck (2004), "Number of sellers, average prices, and price dispersion", *International Journal of Industrial Organization* 22(8-9), 1041-1066.

[8] Borenstein, Severin and Nancy Rose (1994), "Competition and Price Dispersion in the U.S. Airline Industry," *Journal of Political Economy* 102, 653-683.

[9] Burdett, Kenneth and Kenneth L. Judd (1983), "Equilibrium Price Dispersion," *Econometrica* 51, 955-70.

[10] Chandra, Ambarish and Mariano Tapatta (2011), "Consumer Search and Dynamic Price Dispersion: An Application to Gasoline Markets," *Rand Journal of Economics* Vol.42, No.4, Winter 2011.

[11] Chernozhukov, Victor (2005), “Extremal Quantile Regression,” *Annals of Statistics* 33(2), 806-839.

[12] Dana, James D. (1994), “Learning in an Equilibrium Search Model,” *International Economic Review* 35(3), 745-771.

[13] Janssen, Maarten and José L. Moraga-González (2004), “Strategic Pricing, Consumer Search and the Number of Firms”, *Review of Economic Studies* 71(4), 1089-1118.

[14] Galeotti, Andrea (2010), “Searching, Talking and Pricing”, *International Economic Review*, 51(4), 1159-1174.

[15] Gerardi, Kristopher S. and Adam Hale Shapiro (2009), “Does Competition Reduce Price Dispersion: New Evidence from the Airline Industry”, *Journal of Political Economy* 117(1), 1-37.

[16] Haynes, Michelle and S. Thompson (2008), “Price, Price Dispersion and Number of Sellers at a Low Entry Cost Shopbo”, *International Journal of Industrial Organisation*, 26, 459-72.

[17] Hortaçsu, Ali and Chad Syverson (2004), “Product Differentiation, Search Costs and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds”, *Quarterly Journal of Economics* 119, 403-456.

[18] Hosken, Daniel S., Robert S. McMillan and Christopher T. Taylor (2008), “Retail Gasoline Pricing: What Do We Know?”, *International Journal of Industrial Organization* 26(6), 1425-1436.

[19] Houde, Jean-François (2012), “Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline,” *American Economic Review* 102-5, 2147-2182.

[20] Lach, Saul (2002), “Existence and Persistence of Price Dispersion: an Empirical Analysis”, *Review of Economics and Statistics*, August, 433-444.

[21] Lach Saul and José L. Moraga-González (2009), “Asymmetric Price Effects of Competition”, CEPR Discussion Paper 7319.

[22] Lewis, Matthew (2008), “Price Dispersion and Competition with Differentiated Sellers”, *Journal of Industrial Economics* 56 (3), 654-678.

[23] Milgrom, Paul R. (1981), "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics* 12(2), 380-391.

[24] Moldovanu, Benny and Aner Sela (2001), "The Optimal Allocation of Prizes in Contests," *American Economic Review*, 91(3), 542-558.

[25] Morgan, John, Henrik Orzen and Martin Sefton (2006), "An experimental study of price dispersion," *Games and Economic Behavior*, 54, 134–158.

[26] Nermuth, Manfred, Giacomo Pasini, Paolo Pin and Simon Weidenholzer (2009), "Price Dispersion, Search Externalities, and the Digital Divide", mimeo.

[27] NMa (2006): Benzinescan 2005/2006, De Nederlandse Mededingingsautoriteit.

[28] Pennerstorfer, Dieter, Philipp Schmidt-Dengler, Nicolas Schutz, Christoph Weiss and Biliana Yontcheva (2014), "Information and Price Dispersion: Evidence from Retail Gasoline," unpublished manuscript.

[29] Rosenthal, Robert W. (1980), "A Model in Which an Increase in the Number of Sellers Leads to a Higher Price", *Econometrica* 48, 1575–1579.

[30] Soetevent, Adriaan R., Marco A. Haan, and Pim Heijnen (2008), "Do Auctions and Forced Divestitures Increase Competition?", Tinbergen Institute Discussion Paper TI 2008-117/1, The Netherlands.

[31] Varian, Hal (1980), "A Model of Sales," *American Economic Review*, 70, 651-659.

[32] Waterson, Michael (2003), "The Role of Consumers in Competition and Competition Policy," *International Journal of Industrial Organization* 21, 129-50, 2003.

[33] Wildenbeest, Matthijs R. (2011), An empirical model of search with vertically differentiated products, *Rand Journal of Economics*, forthcoming.

[34] Wu, Ximing and Jeffrey M. Perloff (2007), "Tax Incidence Varies Across the Price Distribution," *Economics Letters* 96, 93-99.

Appendix A: Simulations

These are the steps we take to simulate the expected price paid (expected minimum price from a sample of s prices):

- We generate "residual prices" for each station defined by the OLS residuals from regressing price on day-of-the-week dummies and 14 lags of the spot price of gasoline *separately* for each gas station.
- These residual prices represent gasoline prices net of aggregate fluctuations due to wholesale price variation, within-week fluctuations, and store-specific level effects.
- We can therefore treat the residual price vector for each of the N_i firms in market i as coming from the same (time and store invariant) distribution. We now have a sample of (residual) prices for each i.
- We draw s residual prices randomly and without replacement from local market i and store the minium price. This minimum price is the price paid by a consumer observing s prices.
- Repeat this 1000 times and take the average of the 1000 stored minimum prices. This average is our estimate of the expected price paid by a consumer observing s prices. Call this $\hat{y}_{is}$.
- For each local market i, repeat this for each value of $s = 1, \ldots, 4$. We therefore have $\hat{y}_{is}$ for each $s = 1, 2, \ldots, 4$.
- Repeat for each local market i.
- We now have data on $\hat{y}_{is}$, $s = 1, 2, \ldots, 4$ for each market i.

Appendix B: Proofs

Proof of Proposition 2. First we note that

$$q(\alpha_{N+1}(\tau)) - q(\alpha_N(\tau)) = k\left[\frac{\alpha_{N+1}^{(1)}(0)}{\alpha_{N+1}^{(1)}(1-\tau)} - \frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(1-\tau)}\right] = \frac{k}{h(\tau)}\left[\frac{\alpha_N^{(1)}(1-\tau)}{\alpha_N^{(1)}(0)} - \frac{\alpha_{N+1}^{(1)}(1-\tau)}{\alpha_{N+1}^{(1)}(0)}\right] \tag{20}$$

where

$$h(\tau) \equiv \frac{\alpha_{N+1}^{(1)}(0)\alpha_N^{(1)}(0)}{\alpha_{N+1}^{(1)}(1-\tau)\alpha_N^{(1)}(1-\tau)} > 0.$$

(I) Let us set $\tau = 0$ in this expression. It follows that the sign of $q(\alpha_{N+1}(0)) - q(\alpha_N(0))$ equals the sign of

$$-\frac{1}{\alpha_N^{(1)}(1)\alpha_{N+1}^{(1)}(1)}\left[\alpha_{N+1}^{(1)}(1)\alpha_N^{(1)}(0) - \alpha_N^{(1)}(1)\alpha_{N+1}^{(1)}(0)\right] < -\frac{1}{\alpha_{N+1}^{(1)}(1)}\left[\alpha_N^{(1)}(0) - \alpha_{N+1}^{(1)}(0)\right] \leq 0,$$

where the two inequalities follow from Assumption 1. Since (20) is strictly negative when $\tau = 0$, by continuity of the function $q(\alpha_{N+1}(\tau)) - q(\alpha_N(\tau))$ in τ, we conclude that the low percentiles of the price distribution will always decrease.

(II) If condition (7) holds, then it is straightforward to see that all percentiles will fall.

(III) To prove this, we first note that setting $\tau = 1$ in (20) gives 0. Now, let us take the derivative of (20) wrt τ. We get:

$$\frac{\partial}{\partial\tau}(q(\alpha_{N+1}(\tau)) - q(\alpha_N(\tau))) = k\left\{\frac{\alpha_{N+1}^{(1)}(0)\alpha_{N+1}^{(2)}(1-\tau)}{[\alpha_{N+1}^{(1)}(1-\tau)]^2} - \frac{\alpha_N^{(1)}(0)\alpha_N^{(2)}(1-\tau)}{[\alpha_N^{(1)}(1-\tau)]^2}\right\}$$

Setting $\tau = 1$ gives

$$\frac{\partial}{\partial\tau}(q(\alpha_{N+1}(0)) - q(\alpha_N(0))) = k\left[\frac{\alpha_{N+1}^{(2)}(0)}{\alpha_{N+1}^{(1)}(0)} - \frac{\alpha_N^{(2)}(0)}{\alpha_N^{(1)}(0)}\right] \tag{21}$$

When condition (8) holds, this expression is negative. This implies that the difference $q(\alpha_{N+1}(\tau)) - q(\alpha_N(\tau))$ is decreasing in a neighborhood of $\tau = 1$. Since it is zero when $\tau = 1$, by continuity we conclude that $q(\alpha_{N+1}(\tau)) > q(\alpha_N(\tau))$ for sufficiently large percentiles.[37] ■

[37] It could be the case that (21) is zero. In that case, it is straightforward to see that higher order derivatives can be invoked. For example, we can take the second order derivative of (20) and evaluate it at $\tau = 1$, which gives

$$k\left[\frac{\alpha_{N+1}^{(3)}(0)}{\alpha'_{N+1}(0)} - \frac{\alpha_N^{(3)}(0)}{\alpha'_N(0)}\right].$$

If this expression is negative, then the derivative of (20) decreases in a neighborhood of $\tau = 1$. As a result $q(\alpha_{N+1}(\tau)) - q(\alpha_N(\tau))$ must be concave in a neighborhood of $\tau = 1$ and since it itself and its derivative are equal to zero at $\tau = 1$, we conclude $q(\alpha_{N+1}(\tau)) - q(\alpha_N(\tau)) > 0$ in a neighborhood of $\tau = 1$.

Proof of Corollary 1. Using (20), we have that $q(\alpha_{N+1}(\tau)) - q(\alpha_N(\tau)) < 0$ if and only

$$\frac{\alpha_N^{(1)}(1-\tau)}{\alpha_N^{(1)}(0)} - \frac{\alpha_{N+1}^{(1)}(1-\tau)}{\alpha_{N+1}^{(1)}(0)} = \sum_{s=1}^{N} s\frac{\mu_s(N)}{\mu_1(N)}(1-\tau)^{s-1} - \sum_{s=1}^{N+1} s\frac{\mu_s(N+1)}{\mu_1(N+1)}(1-\tau)^{s-1}$$
$$= \sum_{s=1}^{N} s\left[\frac{\mu_s(N)}{\mu_1(N)} - \frac{\mu_s(N+1)}{\mu_1(N+1)}\right](1-\tau)^{s-1} - (N+1)\frac{\mu_{N+1}(N+1)}{\mu_1(N+1)}(1-\tau)^N < 0.$$

From this expression, it is clear that

$$\frac{\alpha_N^{(s)}(0)}{\alpha_N^{(1)}(0)} - \frac{\alpha_{N+1}^{(s)}(0)}{\alpha_{N+1}^{(1)}(0)} = s!\left[\frac{\mu_s(N)}{\mu_1(N)} - \frac{\mu_s(N+1)}{\mu_1(N+1)}\right] \leq 0 \text{ for all } s \text{ suffices.}$$

■

Proof of Corollary 2. Follows straightforwardly from Proposition 2. ■

Proof of Proposition 3. (I) Comparing the expected utility of a consumer who observes s prices when there are N firms and when there are $N+1$ firms gives:

$$CS_{sN+1} - CS_{sN} = k\int_0^1 \left[\frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(\tau^{1/s})} - \frac{\alpha_{N+1}^{(1)}(0)}{\alpha_{N+1}^{(1)}(\tau^{1/s})}\right]d\tau = k\int_0^1 \left[\frac{\alpha_N^{(1)}(0)}{\alpha_N^{(1)}(\tau)} - \frac{\alpha_{N+1}^{(1)}(0)}{\alpha_{N+1}^{(1)}(\tau)}\right] s\tau^{s-1}d\tau \quad (22)$$

When condition (7) holds, the integrand in the above expression is positive and therefore $CS_{sN+1} - CS_{sN} > 0$.

(II) When (8) holds, the integrand of (22) is negative for all $\tau \leq \widetilde{\tau}$ ($\widetilde{\tau}$ as defined in the proof of Proposition 2). To deal with that situation, we follow Janssen and Moraga-González (2004). First, we write

$$CS_{sN+1} - CS_{sN} = k\int_0^1 \frac{\sum_{\ell=1}^{N+1} \ell\left[\frac{\mu_\ell(N+1)}{\mu_1(N+1)} - \frac{\mu_\ell(N)}{\mu_1(N)}\right]\tau^{\frac{\ell-1}{s}}}{\left(\sum_{\ell=1}^{N} \ell\frac{\mu_\ell(N)}{\mu_1(N)}\tau^{\frac{\ell-1}{s}}\right)\left(\sum_{\ell=1}^{N+1} \ell\frac{\mu_\ell(N+1)}{\mu_1(N+1)}\tau^{\frac{\ell-1}{s}}\right)}d\tau \quad (23)$$

and split the integral in (23) as follows:

$$\int_0^1 \frac{\sum_{\ell=1}^{N+1} \ell\left[\frac{\mu_\ell(N+1)}{\mu_1(N+1)} - \frac{\mu_\ell(N)}{\mu_1(N)}\right]\tau^{\frac{\ell-1}{s}}}{\left(\sum_{\ell=1}^{N} \ell\frac{\mu_\ell(N)}{\mu_1(N)}\tau^{\frac{\ell-1}{s}}\right)\left(\sum_{\ell=1}^{N+1} \ell\frac{\mu_\ell(N+1)}{\mu_1(N+1)}\tau^{\frac{\ell-1}{s}}\right)}d\tau$$
$$= -\int_0^{\widetilde{\tau}} \frac{\sum_{\ell=1}^{N+1} \ell\left[\frac{\mu_\ell(N)}{\mu_1(N)} - \frac{\mu_\ell(N+1)}{\mu_1(N+1)}\right]\tau^{\frac{\ell-1}{s}}}{\left(\sum_{\ell=1}^{N} \ell\frac{\mu_\ell(N)}{\mu_1(N)}\tau^{\frac{\ell-1}{s}}\right)\left(\sum_{\ell=1}^{N+1} \ell\frac{\mu_\ell(N+1)}{\mu_1(N+1)}\tau^{\frac{\ell-1}{s}}\right)}d\tau$$
$$+\int_{\widetilde{\tau}}^1 \frac{\sum_{\ell=1}^{N+1} \ell\left[\frac{\mu_\ell(N+1)}{\mu_1(N+1)} - \frac{\mu_\ell(N)}{\mu_1(N)}\right]\tau^{\frac{\ell-1}{s}}}{\left(\sum_{\ell=1}^{N} \ell\frac{\mu_\ell(N)}{\mu_1(N)}\tau^{\frac{\ell-1}{s}}\right)\left(\sum_{\ell=1}^{N+1} \ell\frac{\mu_\ell(N+1)}{\mu_1(N+1)}\tau^{\frac{\ell-1}{s}}\right)}d\tau \quad (24)$$

Notice that the denominator of these integrals increases in τ. Therefore, (24) is lower than

$$
\begin{aligned}
&-\int_0^{\widetilde{\tau}} \frac{\sum_{\ell=1}^{N+1} \ell \left[\frac{\mu_\ell(N)}{\mu_1(N)} - \frac{\mu_\ell(N+1)}{\mu_1(N+1)}\right] \tau^{\frac{\ell-1}{s}}}{\left(\sum_{\ell=1}^{N} \ell \frac{\mu_\ell(N)}{\mu_1(N)} \widetilde{\tau}^{\frac{\ell-1}{s}}\right) \left(\sum_{\ell=1}^{N+1} \ell \frac{\mu_\ell(N+1)}{\mu_1(N+1)} \widetilde{\tau}^{\frac{\ell-1}{s}}\right)} \\
&+\int_{\widetilde{\tau}}^1 \frac{\sum_{\ell=1}^{N+1} \ell \left[\frac{\mu_\ell(N+1)}{\mu_1(N+1)} - \frac{\mu_\ell(N)}{\mu_1(N)}\right] \tau^{\frac{\ell-1}{s}}}{\left(\sum_{\ell=1}^{N} \ell \frac{\mu_\ell(N)}{\mu_1(N)} \widetilde{\tau}^{\frac{\ell-1}{s}}\right) \left(\sum_{\ell=1}^{N+1} \ell \frac{\mu_\ell(N+1)}{\mu_1(N+1)} \widetilde{\tau}^{\frac{\ell-1}{s}}\right)} d\tau \\
&= \frac{\int_0^1 \sum_{\ell=1}^{N+1} \ell \left[\frac{\mu_\ell(N+1)}{\mu_1(N+1)} - \frac{\mu_\ell(N)}{\mu_1(N)}\right] \tau^{\frac{\ell-1}{s}} d\tau}{\left(\sum_{\ell=1}^{N} \ell \frac{\mu_\ell(N)}{\mu_1(N)} \widetilde{\tau}^{\frac{\ell-1}{s}}\right) \left(\sum_{\ell=1}^{N+1} \ell \frac{\mu_\ell(N+1)}{\mu_1(N+1)} \widetilde{\tau}^{\frac{\ell-1}{s}}\right)} \\
&= \frac{\sum_{\ell=1}^{N+1} \ell \left[\frac{\mu_\ell(N+1)}{\mu_1(N+1)} - \frac{\mu_\ell(N)}{\mu_1(N)}\right] \int_0^1 \tau^{\frac{\ell-1}{s}} d\tau}{\left(\sum_{\ell=1}^{N} \ell \frac{\mu_\ell(N)}{\mu_1(N)} \widetilde{\tau}^{\frac{\ell-1}{s}}\right) \left(\sum_{\ell=1}^{N+1} \ell \frac{\mu_\ell(N+1)}{\mu_1(N+1)} \widetilde{\tau}^{\frac{\ell-1}{s}}\right)} \\
&= \frac{\sum_{\ell=1}^{N+1} \frac{\ell s}{s+\ell-1} \left[\frac{\mu_\ell(N+1)}{\mu_1(N+1)} - \frac{\mu_\ell(N)}{\mu_1(N)}\right]}{\left(\sum_{\ell=1}^{N} \ell \frac{\mu_\ell(N)}{\mu_1(N)} \widetilde{\tau}^{\frac{\ell-1}{s}}\right) \left(\sum_{\ell=1}^{N+1} \ell \frac{\mu_\ell(N+1)}{\mu_1(N+1)} \widetilde{\tau}^{\frac{\ell-1}{s}}\right)} \qquad (25)
\end{aligned}
$$

The sign of (25) is equal to the sign of the numerator. Setting $s = 1$ in the numerator of (25) gives

$$\sum_{\ell=1}^{N+1} \left[\frac{\mu_\ell(N+1)}{\mu_1(N+1)} - \frac{\mu_\ell(N)}{\mu_1(N)}\right] = \frac{1}{\mu_1(N+1)} - \frac{1}{\mu_1(N)}$$

which is equal to zero if the condition $\mu_1(N) = \mu_1(N+1)$ is satisfied. As a result, we conclude that $CS_{sN+1} - CS_{sN} < 0$ for $s = 1$. ■