Adv Syst Sci Appl 2023; **01**:129–151
Published online at https://ijassa.ipu.ru.

Verification of Bounded Rationality Models of Price Dispersion on an Online Marketplace Data

Nikolay Bazenkov[1*], Elena Glamozdina[2], Maria Kuznetsova[2], Marina Sandomirskaia[2]

[1] *V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russian Federation*
[2] *HSE University, Moscow, Russian Federation*

Abstract: We explore how well the models of bounded rationality explain price dispersion observed in Internet price-comparison sites. This paper considers three hypotheses about the strategic origin of price dispersion in homogeneous product online sales. Two of them are ε-equilibrium and quantal-response equilibrium (QRE) in a pure Bertrand setting, they involve the bounded rationality of sellers. The third hypothesis introduces the fraction of loyal consumers into the model and assumes the sellers are rational. These behavioral models earlier were proposed in the literature as possible explanations of high price dispersion in online markets. They were supported by experimental lab data, but not tested on real prices. We test the hypotheses on real data for 30 models of household appliances collected from the largest Russian online marketplace Market.Yandex.ru, which organizes e-competition in the closest to the Bertrand setting way. We found no support for ε-equilibrium hypothesis and only limited support for loyal consumer hypothesis. QRE showed the best performance on the data. For most of the products it predicts central tendency, i.e. the mean and the median, remarkably well. The shape of the observed price distributions is explained accurately enough in comparison with the two other hypotheses and random behavior.

Keywords: price dispersion, Internet markets, marketplace, quantal-response equilibrium, ε-equilibrium, loyal consumers. structural estimations

1. INTRODUCTION

Large price dispersion in Internet sales of homogeneous goods is a steady but still unexplained phenomenon. While theoretical analysts originally expected perfect competition in these markets, data-based studies report significant price variations which persist over time [2, 16, 22]. We focus here on the dispersion in posted price, which is higher than that in transacted prices, i.e. the prices at which the real transactions occur. Dispersion in transacted price is a related but a distinct problem which requires access to transactions data which are usually unavailable. So, it is hard to estimate its degree [15, 34].

Below we will discuss several widely used approaches which explain the origin of price dispersion from different points of view. The common for most studies is that sophisticated models generally exist separately from data analysis, which checks tendencies and significant factors without structural estimation of the model. To fill this gap, we take three theoretical models included bounded rationality of market agents and fit them to real price data from a price-comparison site. We show that, in spite of the plausible assumptions behind all three models and their comparable success in prediction of experimental pricing behavior, one of them has much more explanation power than two others.

*Corresponding author: n.bazenkov@yandex.ru

The classical market theory argues that heterogeneous consumer characteristics, in particular, different search costs or information availability, makes a great influence on price setting and price dispersion. This line of research is inspired by the theory of consumer's search behavior [32] and variation in strategies caused by varying costs and access to the complete list of offers [29, 31, 33]. The accurate estimation of the search factor requires consumer-level data, which are generally not available to researchers. Using only price distribution data [20] admits a lot of assumptions on the structure of search costs, such that the analysis of search costs structure for similar products may be inconsistent. Moreover, price-comparison sites are organized in order to reduce consumers' search costs by allowing buyers to order the offers from the cheapest to the most expensive, also with the restrictions on the rating of sellers or on the particular subset of e-shops. They intensify competitive pricing pressure between firms [30]. Under this pressure, firms apply mixed strategies in order to prevent effective learning for consumers [1, 23].

The strategic competitive behavior of e-shops is an equally important source of observed price diversity, so the research of sellers price-setting policies is essential. The common belief [4, 5, 28] is that the large price dispersion must be an equilibrium phenomenon, but the mechanism is unclear. Different theoretical models may produce different equilibrium mixed strategies, so the problem of their selection on data arises.

Some papers analyze markets with the mechanism of simultaneous competing auctions [18, 19]. Together with the consideration of switching problem, as it is costly for consumer to move from one auction at a particular seller to another, authors highlight the influence of additional guarantees on online price dispersion. However, complicated mechanisms poorly fit markets where firms are allowed only to set the price for their good.

In this paper, the models rely on static Bertrand framework which appears a reasonable approximation of a price-comparison platform where firms announce the price for an item. This is caused by the high price-sensitivity of online demand [11], and the fact that most consumers consider the low price as the main factor of purchase, according to numerous consumers' surveys [12, 36, 37]. The static setting is a more questionable modelling assumption, nevertheless, it is widely used in literature, see the survey [4], and the papers comparing multi-channel sales [7, 13, 35], and online advertising [3]. Technically, building a fully dynamic model is complicated by the randomly varying number of active firms at each moment.

Pure Bertrand competition itself can't generate price dispersion, but adding bounded rationality on the firm or on the consumer sides [14] can generate dispersed prices in equilibrium. How these equilibrium distributions correspond to real-world prices is an open question which is addressed in our research. We investigate three hypotheses about the behavior of online sellers. The models underlying these hypotheses were theoretically developed in [2], and [27]. Here they are tested on data from the price-comparison site Market.Yandex.ru. The dataset includes all prices for 30 popular models of household appliances available from July to October 2015 [6]. The prices at every moment show a large dispersion according to the three measures: the coefficient of variation, the relative price range, and the gap between the two minimal prices. We investigate whether the features of the price distributions observed in the data may be explained by the considered theoretical models.

Two hypotheses based on assumptions of bounded rationality of firms are ε-equilibrium and quantal-response equilibrium (QRE) [2]. As indicated in the original paper [2], similar patterns of price dispersion arise both in Internet data and lab experiments, where "unobserved heterogeneities" are excluded by design. However, testing with experimental data does not allow them to conclude that the market behavior is also explained by these bounded rationality models. Both models better work for duopolies, than for triopoly and quadropoly with a large percent of statistical tests rejection, but still better than random or completely rational behavior. Thus, the experiments do not allow to select the best fitted model. Preliminary suggestions on their applicability to Internet markets also looks like the task for this study.

The alternative model concerns the limited rationality of consumers and introduces the fraction of loyal consumers into the classical Bertrand setting [27]. These consumers pick a random seller, not the one who posted the minimal price. In this case, high price dispersion is observed under standard Nash behavior. Though strong experimental support was obtained to the main hypothesis, the question of how this prediction approximates real market behavior remained open and some systematic discrepancies were detected. This model lies close to [10] where some consumers stick to a randomly chosen seller while others search sequentially with an optimal reservation price. While we tested only the model of [27], the intuition is that similar results would be observed for [10]. Though in [27] experiments were designed such that participants knew the share of loyal buyers, in the real data analysis we treat this share as a hidden parameter that is to be estimated. So our estimation strategy is similar with the two hypotheses above, and it is natural to consider all three models in comparison with each other.

In this paper the three hypotheses are reformulated in a form suitable for fitting to the dataset. Applying a similar strategy to [2] we estimated the hidden parameters for each model. Then the goodness of fit is evaluated by three statistical tests. We examined whether each model could produce the mean, the median and the empirical probability distribution for the observed data.

This study demonstrates that the QRE hypothesis survived the tests. Its predictions about the mean and the median are valid, and the prediction about the shape of price distribution is accurate for 50% of the observed time moments. That is much better than the uniform distribution and the two others. The model with loyal consumers performed significantly worse than QRE, but better than pure randomness. There are a non-negligible fraction of days when the observed pricing pattern is similar to those predicted by the model. The ε-equilibrium failed dramatically to explain the observed data despite the reasonable underlying assumption.

To the best of our knowledge this is the first attempt to fit the bounded rationality models to online pricing data. The paper reveals the structure of daily price distributions in homogeneous product online markets. Because of the observed diversity in daily price distributions, the unique theoretical concept doesn't allow to cover all of them. The most frequent situation with the single peak close to the minimal price can be explained by QRE, which is associated with a natural learning process under the changing market conditions. The less frequently observed two-peaked distribution is partly explained by the model with heterogeneous consumers, but this line needs further development in order to select among several close models with potentially non-equal division of loyals [3].

The paper is organized as follows. Section 2 provides a brief description of the dataset from Market.Yandex.ru. Section 3 reproduces the theoretical concepts for further estimations and presents the analytical formulae for price distributions. Section 4 describes the techniques applied, while Section 5 contains the results of the structural estimations of the parameters for each hypothesis. In the last section the quality of theoretical models and their limitations are discussed.

2. ONLINE PRICES: EMPIRICAL OBSERVATIONS

2.1. Data

For the dataset, prices from the online marketplace Market.Yandex.ru were collected during several months from July, 24 to October, 20, 2015 [6]. Market.Yandex.ru is the most popular platform for online shopping in Russia. More than 40% of population and more than 50% of Internet users in Moscow make online purchases[†].

The site is organized in the intuitive way (see Figure 2.1) and free of charge for potential buyers. It is useful both for choosing a specific product by learning their detailed characteristics and for selecting a particular seller of this good. The option of ordering the prices from lowest to highest is available in one click. In addition, it is immediately easy to account for the rating of the e-shop and some delivery conditions. We believe that the rules of a price aggregator makes the structure of the market close to the perfect competition.

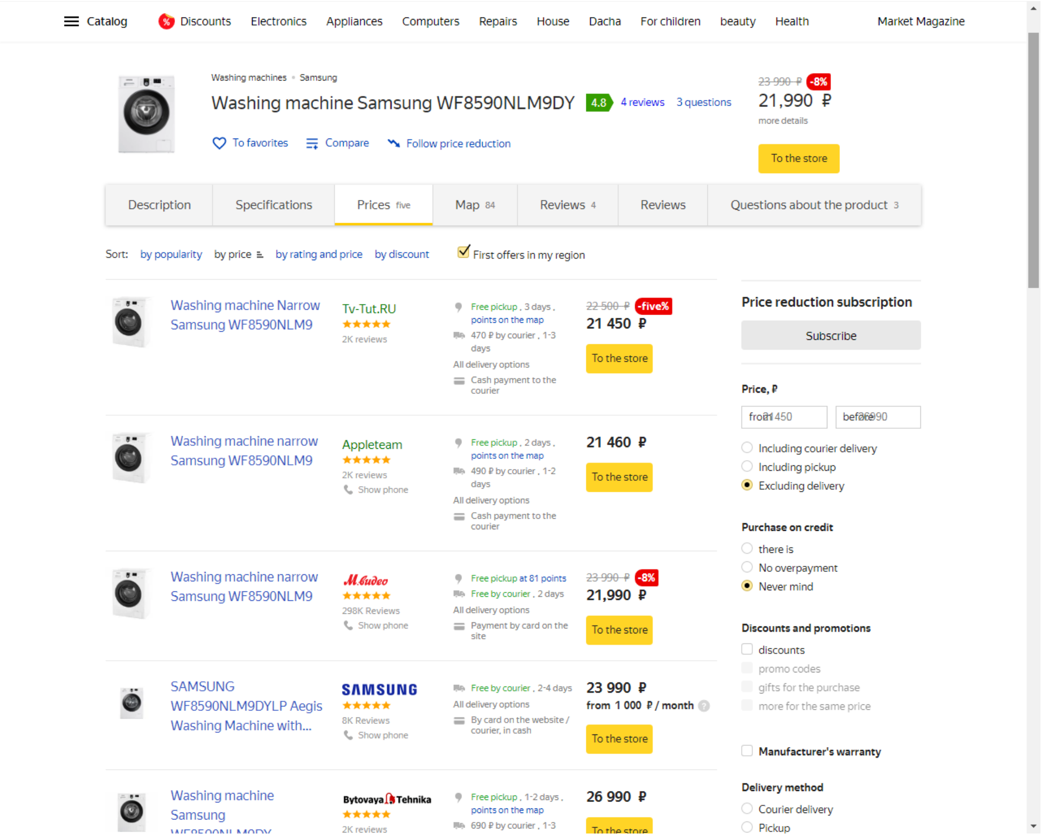

Fig. 2.1. View of Market.Yandex.ru description of the availablle sellers ordered by price for a selected good

We fixed a list of 30 the most popular goods in 5 subcategories of household appliances, i.e. fridges, cooker hoods, warm ovens, dishwashers, cooktops, and washing-machines. The whole set of actual prices was downloaded 4 times per day for e-shops in the Moscow region. Household appliances are in the top-5 categories with the largest share of online sales. For this category the fraction of consumers who prefers to purchase on Market.Yandex.ru exceeds 30%.

Not every e-shop specializing in household appliances sells every product from our sample, and moreover, even if it sells the specific product, the e-shop may not sell it at every moment. This gives unbalanced panel data with 30 goods, 258 sellers and 324 moments containing 463502 unique price offers.

This dataset is representative in several aspects. First, one may be sure that all the offers presented on Yandex.Market.ru are available for the announced price because the platform

[†]Here and further the survey characteristics of industry are based on [12], [36]. In 2021 these shares have obviously become greater.

checks it regularly and sanctions the violators. Second, the metropolitan online market is highly localized: 90% of respondents prefer to buy locally, in Moscow, [36]. Third, geolocation inside the city is not important since for the purchase of major household appliances most consumers use delivery services. Hence this data though covering a limited subset of products along a relatively short time period, may serve a representative example of how competition is organized in modern online markets.

2.2. Price dispersion

The prices in the dataset demonstrate large dispersion both in the coefficient of variation and in the relative price range at each moment. The aggregate description of our data is presented in Table 2.1. The following measures of price dispersion are used. The coefficient of variation is the standard deviation divided by the average. The relative price range is the difference between the maximal and minimal prices divided by the minimal price. The gap is the difference between the two minimal prices divided by the minimal price.

According to these measures, the level of price dispersion we observe for Russian data in 2015 is similar to those reported for USA data in 2001 [2]. Direct comparison to [16] is difficult since they don't report any measures of price dispersion. In [22] the coefficient of variation (relative standard deviation) is 10% for offline retail data from KNRS dataset. This is very close to our average CV which equals 10.6%. We may conclude that general characteristics of our data are similar to those observed in other countries and other samples of goods.

Table 2.1. Average characteristics of products and price dispersion for the whole period

Product	Coeff of var %	Relative price range %	Price gap %	Average price RUB	Lowest price RUB	#
Cata Ceres 600 Negra	22.6	143.6	4.0	19443	12431	36
Electrolux EHH 6340 FOK	15.1	73	2.7	24524	19949	50
Gorenje BO 73 CLI	13.7	70.1	1.4	27435	22947	47
Gorenje gw 65 cl i	13.5	69.3	2.1	17900	14836	28
Electrolux EZB 53400 AX	13.5	51.2	5.2	19944	15750	18
Hansa OSC 511 WH	13.4	79.4	8.4	2813	2124	47
Hotpoint-Ariston 7HPC 640	13	80.5	2.2	11876	9833	67
Krona Kamilla 600	12.7	76.7	13.2	7546	5348	30
Atlant XM 4008-022	12.5	101.1	0.7	14281	12698	59
Hansa ZIM 436 EH	12.2	52.7	1.3	20252	17542	31
Hansa BOEI62000015	11.7	85.7	0.6	16744	14442	68
Electrolux EWS 1052 NDU	11.6	57.5	2.5	19539	16906	31
Hansa BHI68300	11.6	55.1	1.4	16494	14035	45
Bosch DHL 545 S	11.3	65.7	3.3	9220	7706	36
Elikor integra 60n-400	11.1	64.6	2.1	3778	3125	46
Bosch PIC645F17E	10.8	58.4	0.3	28000	23686	78
Indesit BIA 18	10.3	60.4	1.4	21264	17741	71
Bosch HBG43T450	10	60.3	2.7	30429	25717	51
Samsung WF8590NMW9	9.1	78.2	0.7	19111	17217	75
Bosch ActiveWater SPV40E10RU	9.1	47.6	0.8	24489	21671	56
Siemens SR 66T090	8.6	58.3	3.2	43908	37238	68
Hotpoint-Ariston FTR 850	8.5	52.2	1.4	21085	18539	65
Bosch WLG 20061 OE	8.1	54.2	3.0	20417	17506	66
Ariston LSTB 4B00 RU	7.9	43.2	2.2	17519	15710	34
Bosch KGS 39XW20 R	7.8	44.8	1.4	34584	30170	52
Indesit wiun 81 (csi) F053525	7.6	26.7	4.4	13340	12102	9
Candy CDCF 6-07	7	36.4	2.6	13619	12284	25
LG F-1096SD3	6.2	39.6	0.5	21768	19605	67
LG GAB409SVQA	5	28.8	1.6	32571	29610	44
Liebherr SBSesf 7212	4.3	25.1	4.8	124142	111186	29

2.3. Sellers characteristics

For each seller we calculated its average normalized price deviation, i.e. how much it deviates from the average price across all goods and all time moments. Figure 2.2 shows the distribution of average normalized price deviations of the sellers. The distribution is unimodal and slightly skewed to the right. A common assumption is that the observed price dispersion may be to some extent explained by sellers heterogeneity in the quality of service. Below we test this assumption for the data.

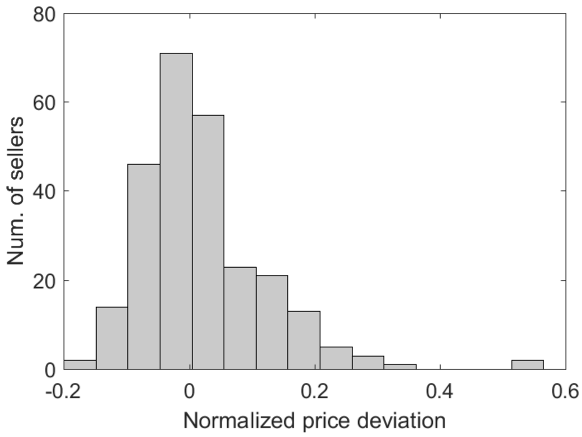

Fig. 2.2. Distribution of sellers normalized price deviations

Yandex.Market platform contains sellers ratings which are numbers from 0 to 5 based on the buyers responses to their experience with the given seller. So, the rating is a long-term reputation which aggregates all information such as delivery conditions, customer support, site reliability etc. The intuition that buyers prefer sellers with good reputation and therefore higher rated sellers will post higher prices doesn't hold for our data. We found no correlation between rating and the average price deviation of the sellers. Figure 2.3a show the relationship between the rating and the average price deviation.

We also used the number of goods a seller offers as a proxy for its size and market share. The intuition is that large shops with diverse assortment are more recognizable, attract more loyal customers and thus may post higher prices. Here we actually observed the opposite tendency. Larger sellers less deviate from the average price than the "small" ones. The correlation between the number of goods available at a seller and its price deviation is -0.24 with $p = 8.2 \times 10^{-5}$. Figure 2.3b shows the relationship between the number of goods available at the seller and the average price deviation.

Another factor which is believed to influence price dispersion is the number of active sellers present in the market. Our data contains dynamic prices and the number of sellers varies greatly across the period. Figure 2.4 shows the distribution of n^t, the number of active sellers observed for the moment t. In the sample at least two firms are present on the market at every moment.

Table 2.2 illustrates how the average price and the dispersion measures correlate with n^t. The most certain tendency is the average price decreasing as n^t rises for 21 of 30 products.

2.4. Shape of the normalized price distributions

Here we describe the general features of the prices in Yandex.Market platform. A single price observation is p_{ik}^t which is the price posted by shop i for product k at the time moment $t \in \{1, ..., T\}$. Data for the given product is a set of price vectors $p^1, \ldots, p^T$ where each vector

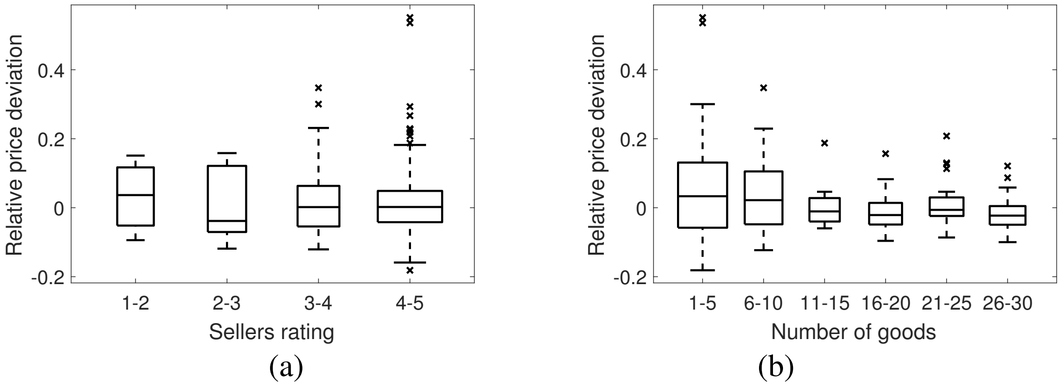

Fig. 2.3. Price deviation relationship with sellers rating and assortment. The widths of the boxes are proportional to the number of sellers in a given category

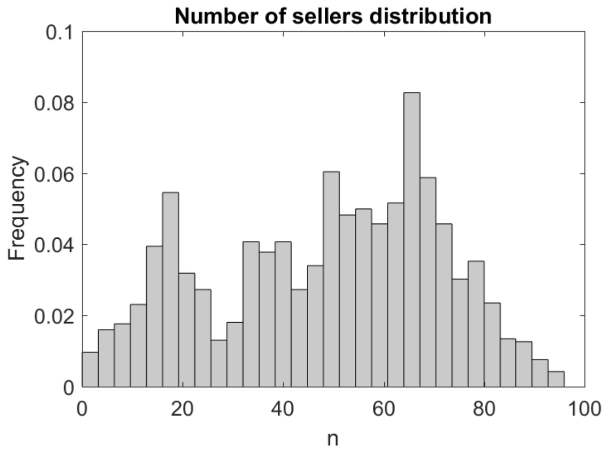

Fig. 2.4. Distribution of the instant number of sellers. The data are aggregated across all days and all products.

Table 2.2. Correlation of price dispersion measures with the number of sellers

	# of products		
	Positive corr.	Negative corr.	Not significant
Average price	5	21	4
Gap	3	17	10
Range	20	7	3
Coeff of var	12	15	3

p^t contains prices observed at moment t. We normalize each price vector to its mean $\overline{p}^t$:

$$\hat{p}^t_{ik} = \frac{p^t_{ik}}{\overline{p}^t}. \tag{2.1}$$

The normalized prices across all days can be aggregated into a single price distribution for each product. While the shapes of these distributions vary from product to product, they share some key features. Figure 2.5 shows the distribution for one of the products, a cooker hood *Bosch DHL 545 S*. The distribution is unimodal, similar to normal, but skewed to the right and leptokurtic. This shape is observed with some variations for all of the examined products. The right tail and the central peak may be more or less prominent, but the general pattern is the same. In [21] similar distributions were observed though they are more symmetric than ours.

We can see that the shape of aggregated price distribution is close to normal. However,

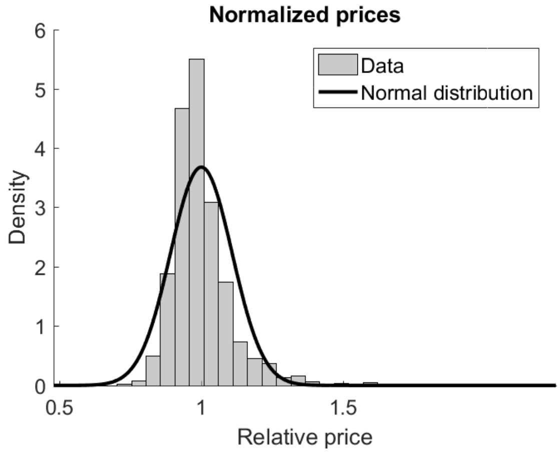

Fig. 2.5. The distribution of normalized prices across all days. The solid curve shows the normal distribution of the same mean and variance.

the distributions for shorter time periods are expected to be be different from the aggregated one. Moreover, they can belong to different classes of distributions, i.e. be one-peaked or two-peaked, for some particular days for the same product. Testing several reasonable theoretical models in order to select one that better fits the data is addressed in the rest of the paper.

3. DAILY PRICE DISTRIBUTION: BOUNDED RATIONALITY MODELS

We examine the behavioral models developed in [2] and [27]. This section briefly introduces the theoretical framework, the intuition behind it and the price distributions under every hypothesis for fitting the models to the data. More formalities are in Appendix A.

3.1. Price competition with homogeneous product

Let us start with the pure homogeneous Bertrand setting. The models are based on the classical Bertrand price competition where n firms sell the same homogeneous product with costs c. The consumers are rational and always buy from the firms who announced the minimal price. This is in line with the design of the online marketplace where consumers are able get the complete list of prices for a chosen product and to sort these prices with a single click. Moreover, we can expect that those consumers who consciously use price-comparison site do this in order to find the best price. Nash equilibrium in this setting yields the Bertrand paradox with zero profits for all firms. This clearly contradicts the observations of real online markets. Seeking a plausible theoretical background of online sellers behavior, the authors in [2] enriched the setting by two models of bounded rationality.

ε-equilibrium The ε-equilibrium implicates the minimal level of extra payoff which is expected to be reached after a deviation. This also can be interpreted as the decision maker being insensitive to small price differences or having limited cognitive abilities. Another logic beyond this concept involves the possible costly price changes and as a result of the lack of motivation for small extra gains from frequent price adjustments.

For the profit function with non-elastic demand normalized to 1, the ε-equilibrium price distribution for $\varepsilon \in (0, [n/(n-1)]^{n-1}2^{-n}(p^m - c))$ is given by

$$F_i^\varepsilon(p) = \begin{cases} 0, & p < c + \theta; \\ 1 - \left(\frac{\theta}{p-c}\right)^{\frac{1}{n-1}}, & p \in [c+\theta, p^m); \\ 1, & p \geq p^m, \end{cases} \tag{3.2}$$

where n is the number of active firms, p^m is the monopoly price, c are costs which are the same for all firms, and

$$\theta = \left[\varepsilon^{n-1}\left(\frac{n}{n-1}\right)^{n-1}(p^m - c)\right]^{1/n} \tag{3.3}$$

Prices generated by this model are concentrated around the two points: θ and p^m. Figure 3.6 the bimodal shape with very low density between the two peaks. The equation for CDF (3.2) implies that the CDF is discontinuous in p^m, so the density in p^m is represented by Dirac delta function. This is shown as the vertical line surmounted by an arrow in the figure. As n grows, θ approaches zero, and for $n \geq 8$ almost all prices fall near the monopoly price p^m or near the minimal price c.

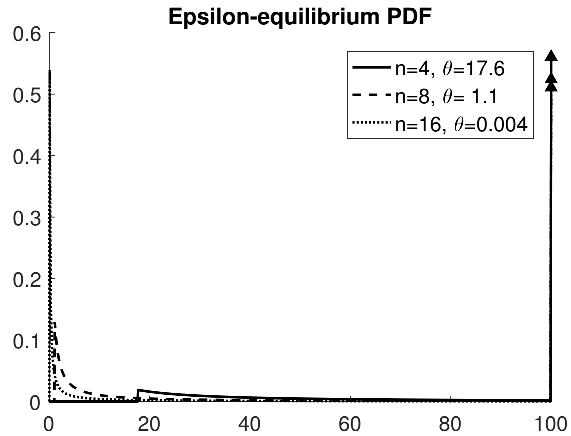

Fig. 3.6. Price distribution for ε-equilibrium. ε is set as 0.5 of the upper bound for the given n.

ε-hypothesis: The firms selling identical products on the Internet play their ε-equilibrium price strategies in the Bertrand model with homogeneous products.

Quantal-response equilibrium Quantal-response equilibrium was proposed in [26]. The idea is that a player randomizes over all available actions with a tendency to choose actions proportionally to their expected utility. In [9] the concept is interpreted as an agent maximizing the "subconscious" utility function which is not completely known to her. It is proved that fictious play learning dynamics converges to QRE [9]. In the context of online retail, this means that a seller can only approximately anticipate what profit it will gain from a particular decision, but after some learning process the policy will favor more profitable actions in a proportion predicted by QRE.

In QRE the probability of setting a particular price depends on the expectations of how this price will influence the profit level under some fixed price distributions of competitors. The motivation for this concept stems from preference shocks [25], or decision errors [24].

For the decision rule power function with parameter $\lambda \in \left[0, \frac{1}{n-1}\right)$, the symmetric QRE is provided by the following price distribution:

$$F_i^Q(p) = \begin{cases} 0, & p < c; \\ 1 - \left(1 - \left(\frac{p-c}{p^m-c}\right)^{(1+\lambda)}\right)^{\frac{1}{1+\lambda-n\lambda}}, & p \in [c, p^m); \\ 1, & p \geq p^m. \end{cases} \tag{3.4}$$

The parameter λ expresses the degree of irrationality in a specific sense. The two limit cases correspond to Nash equilibrium ($\lambda \to \frac{1}{n-1}$) and random behavior ($\lambda \to 0$). The intermediate values of λ generate the set of price distribution functions with a larger mass of firms setting low prices in comparison to ε-equilibrium.

According to this model the price distribution should be unimodal and right skewed, as shown in Figure 3.7. The distribution spans the whole price range $[c, p^m]$. The peak is close to the minimal price. As the players behavior tend from strict rationality to randomness, the distribution becomes flatter, tending to uniform. In contrast to ε-equilibrium, the probability function is smooth and shows more flexibility. The number of sellers n slightly affect, but the effect is much weaker than for ε-equilibrium.

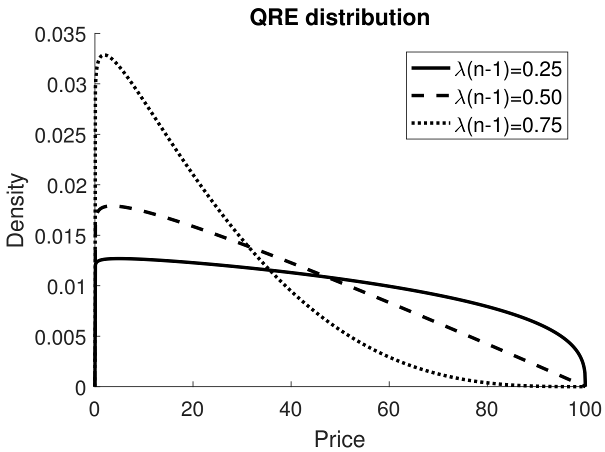

Fig. 3.7. Price distribution for quantal-response equilibrium. The number of sellers $n = 8$, $lambda(n-1)$ is the "degree of rationality": 0 is complete randomness, 1 is strict Nash rationality.

Q**-hypothesis:** The firms selling identical products on the Internet set their prices according to the QRE price distribution in the Bertrand model with homogeneous products.

3.2. Price competition with a share of loyal consumers

An alternative model for explaining the behavior of online competitors is based on consumer heterogeneity in access to the complete list of offers. In particular, this line relates to modelling multi-channel competition via both online sites and bricks-and-mortar stores, in opposite to considering online and offline competitions as independent settings [8].

The paper [27] presents the clearinghouse model which modifies Bertrand price competition by introducing a share of captive consumers. Assume that the fraction α of all consumers (the total mass is normalized to 1) use a price aggregator and choose the firm with the minimal price. The rest of the consumers are not informed about the whole range of firms: they are loyal to a certain firm from the list and buy from it with the reservation price normalized to 1. Here we do not specialize the reasons for this loyalty, which can be the result of prohibitively high search costs or advertising efforts of the firm [3]. As above, costs are equal to c for all firms. The utility of firm i is combined from two sources: non-competitive revenue coming from loyal consumers and a possible gain for the winner of the Bertrand competition. The (Nash) equilibrium price distribution is given by

$$F_i^{\alpha}(p) = \begin{cases} 0, & p < \underline{p}; \\ 1 - \left[\frac{1-\alpha}{\alpha n}\frac{p^m - p}{p - c}\right]^{\frac{1}{n-1}}, & p \in [\underline{p}, p^m); \\ 1, & p \geq p^m, \end{cases} \tag{3.5}$$

where $\underline{p} = \frac{p^m(1-\alpha)+\alpha nc}{1-\alpha+\alpha n}$ is the minimal price for an active firm in the market.

While the assumptions under this model are different from those for ε-equilibrium, the shape of the distribution is similar. The probability density is U-shaped with the first peak at $\underline{p}$, approaching c as α grows, and the second peak at p^m. Figure 3.8 shows that as the fraction of rational consumers α grows, more prices are concentrated around the left bound, though the right peak does not diminish completely even for very high α.

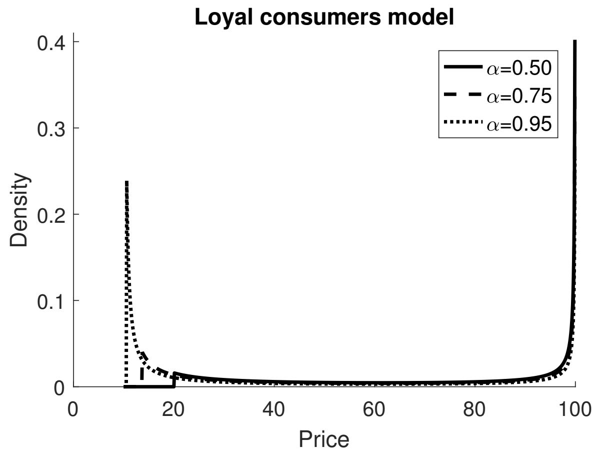

Fig. 3.8. Price distribution for loyal consumers equilibrium. The number of sellers $n = 8$, α is the fraction of rational consumers

α**-hypothesis:** The firms selling identical products on the Internet compete in price for the share of informed consumers within the Bertrand model with homogeneous product and simultaneously get extra revenue from the consumers loyal to them.

4. ESTIMATION TECHNIQUE

Here we examine the shape of the price distributions observed at any particular moment of time. We address the following questions:

1. Does any of the three hypotheses from Section 3 provide a valid explanation for the shape of the observed price distributions;
2. What probability distribution is the best parametric model for the observed price data.

To answer the first question we apply the technique similar to [2]. For each of the three hypotheses, we statistically tested how well it predicts the mean, the median, and the histogram of the observed price distributions.

In the analyzed behavioral models the number of sellers n greatly influences the shape of the price distribution, so the analysis was performed separately for the different numbers of sellers which is exogenous under the complete access to information about the list of competitors. The data were grouped such that each group contains only price vectors of the given length n and the total number of prices in each group was at least $N = 25$. Then the parameters were estimated separately for each group.

Each behavioral model contains three parameters. The monopoly price p^m, the cost c, and the shape parameter: λ, θ, and α. In the previous work [2] the monopoly price p^m and the cost c were fixed by the experimental design. Here all three parameters are estimated by minimizing the sum of squared errors between the empirical and the theoretical cumulative distribution functions, following the same approach as [2].

After the parameters had been estimated for the given price group, we evaluate the observed and theoretical distributions by three tests. T-test was used to compare the observed mean and the mean predicted by each theoretical model. The sign test was used to compare the observed and the predicted medians.

The goodness of fit was evaluated by the chi-square test. Five bins were used, chosen such that the expected frequencies were approximately equal for each bin. The significance level was chosen as $p = 0.05$ for each test. We calculated the fraction of all time moments when the hypothesis was not rejected. This fraction estimates how well the given behavioral model explains the observed data *post factum*. The next section covers the results of these estimations.

5. RESULTS

The complete results are summarized in Tables 6.4, 6.5, and 6.6, provided in Appendix. The percent of successful tests are shown in Table 5.3. The “Random” column corresponds to uniformly distributed prices. A detailed discussion for each hypothesis is presented below.

Table 5.3. The fraction of time when the hypotheses are not rejected, the significance level is $p = 0.05$

Test	$Q-$hypothesis	$\varepsilon-$hypothesis	$\alpha-$hypothesis	Random
Mean	98%	20%	33%	14%
Median	92%	12%	43%	15%
Distribution	49%	0.1%	3%	0.4%

5.1. Success of Q-hypothesis

The results of estimation of Q-hypothesis are the most optimistic. The mean and median tests fail to reject QRE in almost 100% of the tests. This is caused by the fact that most of the observed price distributions are unimodal with positive skewness and this shape is in general similar to that induced by the QRE hypothesis. One of the successful fits is presented in Figure 5.9.

The goodness-of-fit chi-square test rejects the null hypothesis for a half of the observed moments. The best result is for the cooktop Hotpoint-Ariston 7HPC 640, where QRE hypothesis passed the test in 91% of the time moments. For a half of the products QRE passed the test in more than 55% of time. There is a single product, the fridge Liebherr SBSesf 7212, with 0% of QRE passed, this may be due to its high price in comparison with other products in our sample, and thus, a specific consumer demand structure not covered by the model assumptions. Figure 5.10 shows the success rate varying from 90% to 0%. The success rate does not depend on the number of sellers.

The fraction of successful estimations convinces us that QRE hypothesis performs much better than the uniform prices and the other two hypotheses. We assume that this result is due to the great flexibility of quantal response equilibrium which covers the broad range of distribution shapes. Therefore, our study shows that QRE can justify very different behavioral patterns, though this model is still falsifiable since it was rejected in 50% of the time. The model may serve as a useful technique to weaken the assumption of strict rationality. Further

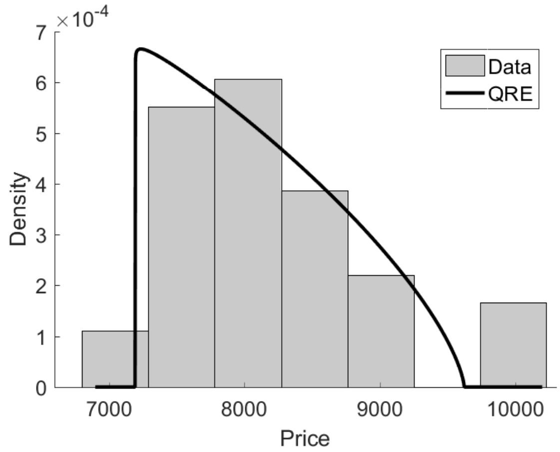

Fig. 5.9. Empirical and the theoretical distributions for the kitchen hood Bosch DHL 545 S. Number of sellers is $n = 37$. The prices are from 2015, August 17 at 1am. Only QRE passed χ^2 test ($\lambda(n-1) = 0.41$, $p_{\chi^2} = 0.18$).

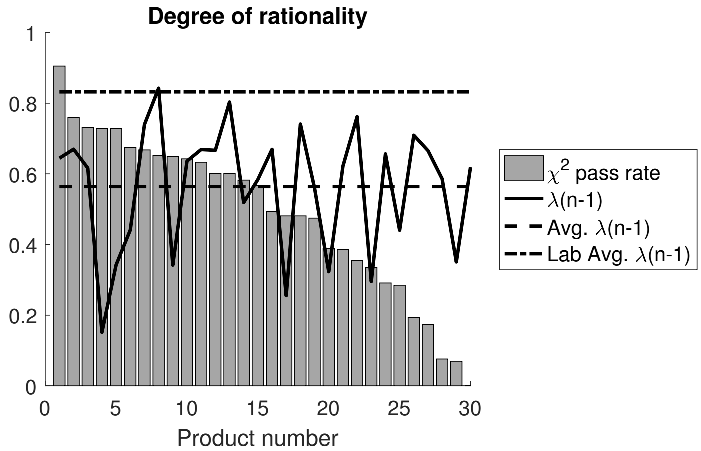

Fig. 5.10. Fraction of QRE passed χ^2 test for all products (bars) and the related degree of rationality $\lambda(n-1)$ (solid line), average is 0.56 (dashed line) and the average value for lab experiments in [2] is 0.83

this can be applied in combination with some sophisticated market model involving multi-channel competition or consumer search.

5.2. *Testing of ε- and α-hypotheses*

The estimations of ε- and α-hypotheses are much less successful than QRE. The prediction quality of ε-hypothesis is comparable to the uniform distribution. Even the mean and the median tests provide weak predictions, while the chi-squared test fails dramatically. The examples of successful fits are shown in Figure 5.11 and Figure 5.12.

A possible reason for the poor performance is that both ε- and α-hypotheses predict high concentration of prices near the monopoly price for $n \geq 8$. But this pattern is rarely observed in our data, most of the time the prices lie close to the minimal price, not maximal. Table 2.2 indicates that, for most of the products, the average price correlates negatively with the number of sellers. This tendency also directly contradicts the predictions of both models.

It is to be noted that the model with loyal consumers, α-hypothesis, performs significantly better than Random and ε-hypotheses. The reason is that this model generates the smooth distribution curve which is more flexible than those of ε-equilibrium. Nevertheless, the fraction of successful estimations is considerably lower than that for QRE. It may be the

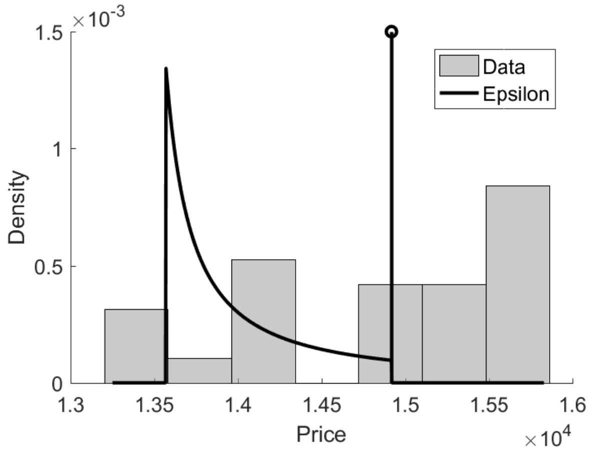

Fig. 5.11. Empirical and the theoretical distributions for the washing machine Indesit wiun 81 (csi) F053525. Number of sellers is $n = 5$. Prices are grouped across 5 time points starting from September 14. Only ε-equilibrium passed χ^2 test: $\theta = 183.8$, $p_{\chi^2} = 0.08$.

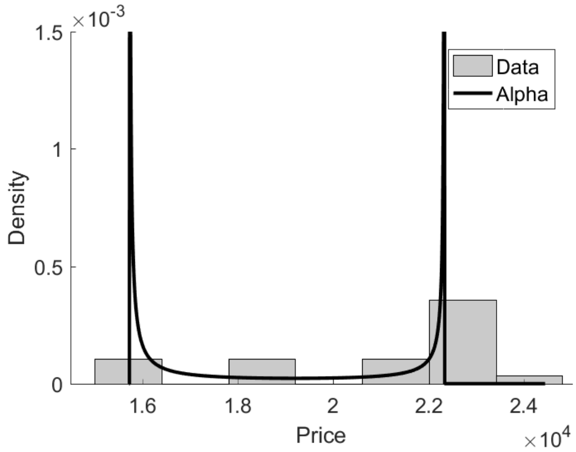

Fig. 5.12. Empirical and the theoretical distributions for the warm oven Electrolux EZB 53400 AX. Number of sellers $n = 20$, prices are grouped from August 19, 1am and 7am. Only α-hypothesis passed χ^2 test: $\alpha = 0.97$, $p_{\chi^2} = 0.13$

case that for some particular situations the behavioral patterns predicted by α-hypothesis do prevail.

The theoretical distributions generated by these two hypotheses have an important distinction: the right peak in ε-hypothesis is an "anomaly", i.e. the price density is discontinuous here and the density just in the left neighbourhood of the monopoly price tends to zero. For α-hypothesis we observe the smooth, U-shaped distribution with positive densities on the both sides of the price range. This corresponds to the determination of each firm to serve either "shoppers" (for firms with the price close to the marginal cost), or "loyals" (local monopolists) at every moment.

6. DISCUSSION

The objective of this study was to verify the models from [2], and [27] on real price data. The results show that, from the three models to explain Internet price dispersion, one performed significantly better than the other two, and one performed not better than random behavior.

The principle difficulty concerns the number of competitors. In the Bertrand model with ε-equilibrium and in the model with loyal consumers the higher the number of sellers, the higher the prices must be concentrated at the monopoly price, while our data shows the opposite tendency. In the experiments, reported in [2], and [27] there were no more than four competitors, which is in line with the model. However, in Internet sales, the number of active firms is generally greater, such that these models are not applicable. Even in the case of a relatively small number of sellers, the predictions under ε-equilibrium are not valid, and the predictions of $\alpha-$hypotheses are rarely successful. These results were not obvious, because the assumptions underlying both models have a reasonable interpretation in the context of online markets.

QRE performs much better in comparison with the two others. It is shown that the central tendency is predicted very accurately by QRE. The fraction of successful distribution fits is 50%, which is surprisingly high. The average value of the predicted degree of rationality $\lambda(n-1)$ is 0.56, closer to 0 than the average 0.83, reported in the experiments [2]. According to the model this means that real firms are *less rational* than the participants of the lab session. This interpretation contradicts the intuition that online markets are the highly competitive environment where sellers employ highly optimized strategies aided by modern software tools.

One of the possible interpretation of this “irrationality” is the following. Online markets are highly stochastic, and in contrast to Lab experiments the firms do not know exactly what price is optimal at this particular moment. Hence they are trying to pick the optimal price through some trial-and-error iterative process, which never converges since the environment evolves permanently in an unpredictable direction. The motivation behind quantal-response equilibrium proposed in [9] reflects this vision. The agents cannot compute the optimal price precisely, but they have more or less correct intuitive assumptions about what actions are more profitable. So, the result that agents in real market rely on randomization more than in Lab experiments is not due to their less rationality, but due to the random nature of the environment.

The relative success of QRE is also explained by the fact that QRE admits a much broader set of possible price distributions. An important note is that the general formulation of QRE is proved to lack falsifiability until additional restrictions are introduced [17]. Here we examined only symmetrical QRE, and this formulation appeared to be sufficiently strict since only half of all price distributions doesn’t contradict this hypothesis according to χ^2.

To summarize, our observation supports the important role of bounded rationality for the adequate modeling of online competition. One more optimistic observation is that the pure Bertrand model occurs to be enough for obtaining successful estimations with relatively high explanatory power.

The question of what model is suitable still remains open, but QRE has proved itself as a flexible model which can be fit successfully to the real-life data. The model with loyal consumers appears to catch at least some features of the observed behavior. However, this model requires the further elaboration.

The combination of both approaches may be a promising direction of future research. Another possible direction is to extend the homogeneous Bertrand competition in a way which is more suitable for online markets, i.e. to introduce the idea of online product differentiation.

ACKNOWLEDGMENTS

Maria Kuznetsova and Marina Sandomirskaia acknowledge the suppport of the HSE University Basic Research Program. Nikolay Bazenkov acknowledges the financial support by Russian Foundation for Basic Research under project #20-07-00190A. The authors gratefully thank the Yandex.Market Customer Support API of commercial services for the data access. The authors are indebted to Jacques-Francois Thisse, Sergei Izmalkov, Alexander Tarasov, Fuad Aleskerov, and Alexis Belianin for fruitful discussions and useful comments. We also thank Ernan Haruvy for the detailed recommendations which allow us to improve the performance of the results significantly.

APPENDIX A. THEORETICAL FRAMEWORK

Bertrand competition

Let firm $i \in \{1, \ldots, n\}$ proposes price p_i simultaneously and independently of all other firms. For simplicity, let $p_i \in [c,\, p^m]$, p^m be the monopoly price, and $\pi(p^m)$ be the monopoly payoff. The expected utility of firm i under the price vector $(p_1, \ldots, p_i, \ldots, p_n)$ is given by

$$\pi_i(p_1, \ldots, p_i, \ldots, p_n) = \begin{cases} \frac{\pi(p_i)}{k}, & p_i = \min\{p_1, \ldots, p^n\}, \\ & k \text{ is the number of firms charging } p_i; \\ 0, & \text{otherwise.} \end{cases} \tag{6.6}$$

For estimation purposes the monopoly profit function with non-elastic demand normalized to 1 is used:

$$\pi(p) = p - c. \tag{6.7}$$

The mixed strategy expansion is considered with a price cumulative distribution vector $\{F_1, \ldots, F_n\}$. The expected payoff of firm i is

$$E\pi_i(F) = \int_{[c,\, p^m]^n} \pi_i(p) dF.$$

ε-equilibrium

Definition 1. Fix $\varepsilon > 0$. A strategy vector F^ε is an ε-equilibrium if for any unilateral deviation F_i' of player i

$$E\pi_i(F_i', F_{-i}^\varepsilon) - E\pi_i(F^\varepsilon) \leq \varepsilon, \quad i = 1, \ldots, n.$$

In ([2]), the solution for the Bertrand model with arbitrary function π was obtained. For the profit function (6.7), the ε-equilibrium price distribution is given by formulae (3.2).

Quantal-response equilibrium

Let T_i be a map from $E\pi_i(p_i, F_{-i})$ into the probability of setting the price less than or equal to p_i.

Definition 2. A strategy vector F^Q is a quantal-response equilibrium if for any p_i

$$F_i^Q(p_i) = T_i(E\pi_i(p_i, F_{-i}^Q)), \quad i = 1, \ldots, n.$$

In ([2]) the decision rule function T was set as the power function with parameter $\lambda \in \left[0, \frac{1}{n-1}\right)$ []

$$T_i(E\pi_i(p, F_{-i})) = \int_c^p \frac{E\pi_i(q, F_{-i})^\lambda}{\int_c^{p^m} E\pi_i(t, F_{-i})^\lambda dt} dq. \tag{6.8}$$

For T_i defined by (6.8), the symmetric QRE is provided by the following price distribution:

$$F_i^Q(p) = 1 - \left(\frac{g(p^m) - g(p)}{g(p^m) - g(c)}\right)^{1/(1+\lambda-n\lambda)}, \quad i = 1, \dots, n, \; p \in [c, p^m],$$

with $g(p) \equiv \int \pi(p)^\lambda dp + K$.

Accounting for (6.7), the final form of the QRE price distribution is (3.4).

Bertrand competition with a share of loyal consumers

The utility of firm i has a view

$$\pi_i(p_1, \dots, p_n) = \begin{cases} \frac{\alpha(p_i - c)}{k} + \frac{(1-\alpha)(p_i - c)}{n}, & p_i = \min\{p_1, \dots, p^n\}, \\ & k \text{ is the number of firms charging } p_i; \\ \frac{(1-\alpha)(p_i - c)}{n}, & \text{otherwise.} \end{cases} \tag{6.9}$$

By analogy with the logic in [27], it is easy to derive formula (3.5). The corresponding lower bound for the fraction α is determined by

$$\underline{\alpha} = \frac{p^m - \underline{p}}{p^m - nc + \underline{p}(n-1)}. \tag{6.10}$$

The search-theoretic model proposed in [10] employs a different set of assumptions, but provides similar U-shaped price distribution. So the results obtained for ε-equilibrium and α-hypotheses may be generalized to this model.

APPENDIX B. ESTIMATION RESULTS

Table 6.4. Tests of equality of Means, Medians, and Distributions: the shares of moments for every product for which Q- and ε-hypotheses are not rejected, the significance level is 0.05

	QRE			$\varepsilon-$equilibrium		
Product	Mean	Median	χ^2	Mean	Median	χ^2
Bosch DHL 545 S	1.00	0.99	0.60	0.12	0.10	0.00
Hotpoint-Ariston FTR 850	1.00	1.00	0.73	0.00	0.00	0.00
Indesit wiun 81 (csi) F053525	0.85	0.93	0.17	0.43	0.39	0.04
Elikor integra 60n-400	1.00	1.00	0.48	0.26	0.13	0.00
Hansa OSC 511 WH	1.00	1.00	0.76	0.28	0.19	0.00
Krona Kamilla 600	0.94	0.35	0.07	0.55	0.25	0.00
Samsung WF8590NMW9	1.00	1.00	0.57	0.13	0.00	0.00
Atlant XM 4008-022	1.00	1.00	0.47	0.00	0.00	0.00
Liebherr SBSesf 7212	0.84	0.06	0.00	0.13	0.22	0.00
Candy CDCF 6-07	1.00	0.97	0.60	0.20	0.20	0.00
Siemens SR 66T090	1.00	0.88	0.19	0.00	0.00	0.00
Bosch ActiveWater SPV40E10RU	1.00	0.91	0.35	0.03	0.02	0.00
Hansa BOEI62000015	1.00	1.00	0.65	0.00	0.00	0.00
Bosch HBG43T450	1.00	1.00	0.39	0.82	0.02	0.00
Indesit BIA 18	1.00	1.00	0.67	0.03	0.00	0.00
Hotpoint-Ariston 7HPC 640	1.00	1.00	0.91	0.03	0.00	0.00
Cata Ceres 600 Negra	1.00	1.00	0.67	0.23	0.31	0.00
Electrolux EHH 6340 FOK	1.00	1.00	0.58	0.30	0.14	0.00
Bosch KGS 39XW20 R	1.00	0.84	0.34	0.26	0.01	0.00
LG GAB409SVQA	1.00	1.00	0.65	0.12	0.12	0.00
Bosch WLG 20061 OE	1.00	0.95	0.49	0.02	0.05	0.00
Hansa ZIM 436 EH	0.99	1.00	0.48	0.28	0.20	0.00
Bosch PIC645F17E	1.00	0.92	0.29	0.35	0.00	0.00
Hansa BHI68300	1.00	1.00	0.73	0.18	0.17	0.00
LG F-1096SD3	1.00	0.98	0.64	0.03	0.00	0.00
Gorenje BO 73 CLI	0.95	0.96	0.28	0.13	0.07	0.00
Gorenje gw 65 cl i	1.00	0.97	0.39	0.13	0.23	0.00
Electrolux EWS 1052 NDU	0.96	1.00	0.63	0.00	0.04	0.00
Electrolux EZB 53400 AX	0.99	0.96	0.08	0.98	0.55	0.00
Ariston LSTB 4B00 RU	1.00	1.00	0.73	0.12	0.16	0.00

Table 6.5. Tests of equality of Means, Medians, and Distributions: the shares of moments for every product for which α-hypothesis and Random behavior are not rejected, the significance level is 0.05

	α-			Random		
Product	Mean	Median	χ^2	Mean	Median	χ^2
Bosch DHL 545 S	0.34	0.48	0.01	0.00	0.03	0.00
Hotpoint-Ariston FTR 850	0.10	0.38	0.00	0.00	0.00	0.00
Indesit wiun 81 (csi) F053525	0.88	0.91	0.12	0.44	0.54	0.02
Elikor integra 60n-400	0.83	0.38	0.01	0.06	0.12	0.00
Hansa OSC 511 WH	0.52	0.69	0.03	0.38	0.43	0.01
Krona Kamilla 600	0.45	0.32	0.07	0.37	0.17	0.00
Samsung WF8590NMW9	0.08	0.43	0.00	0.00	0.01	0.00
Atlant XM 4008-022	0.18	0.28	0.00	0.00	0.00	0.00
Liebherr SBSesf 7212	0.50	0.33	0.03	0.27	0.03	0.00
Candy CDCF 6-07	0.53	0.64	0.06	0.16	0.14	0.01
Siemens SR 66T090	0.12	0.12	0.00	0.01	0.00	0.00
Bosch ActiveWater SPV40E10RU	0.04	0.21	0.00	0.00	0.00	0.00
Hansa BOEI62000015	0.08	0.25	0.00	0.00	0.00	0.00
Bosch HBG43T450	0.04	0.15	0.00	0.04	0.00	0.00
Indesit BIA 18	0.21	0.53	0.00	0.02	0.02	0.00
Hotpoint-Ariston 7HPC 640	0.22	0.47	0.00	0.10	0.09	0.00
Cata Ceres 600 Negra	0.31	0.29	0.00	0.29	0.62	0.00
Electrolux EHH 6340 FOK	0.35	0.41	0.01	0.24	0.21	0.00
Bosch KGS 39XW20 R	0.13	0.21	0.00	0.09	0.11	0.00
LG GAB409SVQA	0.18	0.21	0.00	0.29	0.25	0.02
Bosch WLG 20061 OE	0.12	0.44	0.00	0.01	0.00	0.00
Hansa ZIM 436 EH	0.36	0.56	0.03	0.14	0.12	0.00
Bosch PIC645F17E	0.36	0.54	0.00	0.00	0.06	0.00
Hansa BHI68300	0.30	0.40	0.00	0.09	0.23	0.01
LG F-1096SD3	0.17	0.54	0.00	0.08	0.08	0.00
Gorenje BO 73 CLI	0.24	0.31	0.04	0.00	0.01	0.00
Gorenje gw 65 cl i	0.50	0.47	0.03	0.03	0.09	0.00
Electrolux EWS 1052 NDU	0.50	0.60	0.17	0.13	0.10	0.00
Electrolux EZB 53400 AX	0.76	0.91	0.15	0.80	0.80	0.01
Ariston LSTB 4B00 RU	0.40	0.43	0.05	0.09	0.09	0.03

Table 6.6. Parameters of QRE and α-hypothesis

Product	$\lambda(n-1)$ All days	Success fit days	α All days	Success fit days
Bosch DHL 545 S	0.64	0.61	0.99	1.00
Hotpoint-Ariston FTR 850	0.67	0.70	1.00	-
Indesit wiun 81 (csi) F053525	0.62	0.86	0.84	0.97
Elikor integra 60n-400	0.15	0.17	1.00	1.00
Hansa OSC 511 WH	0.34	0.33	0.99	0.99
Krona Kamilla 600	0.44	0.71	0.88	0.96
Samsung WF8590NMW9	0.74	0.66	1.00	-
Atlant XM 4008-022	0.84	0.78	1.00	-
Liebherr SBSesf 7212	0.34	-	0.87	0.94
Candy CDCF 6-07	0.64	0.64	0.99	0.95
Siemens SR 66T090	0.67	0.73	1.00	-
Bosch ActiveWater SPV40E10RU	0.67	0.71	1.00	-
Hansa BOEI62000015	0.80	0.81	1.00	-
Bosch HBG43T450	0.52	0.59	1.00	-
Indesit BIA 18	0.58	0.60	1.00	-
Hotpoint-Ariston 7HPC 640	0.67	0.67	1.00	-
Cata Ceres 600 Negra	0.26	0.29	1.00	-
Electrolux EHH 6340 FOK	0.74	0.77	1.00	1.00
Bosch KGS 39XW20 R	0.56	0.66	1.00	-
LG GAB409SVQA	0.32	0.28	1.00	-
Bosch WLG 20061 OE	0.62	0.59	1.00	-
Hansa ZIM 436 EH	0.76	0.76	1.00	1.00
Bosch PIC645F17E	0.30	0.37	1.00	-
Hansa BHI68300	0.66	0.66	1.00	-
LG F-1096SD3	0.44	0.37	1.00	-
Gorenje BO 73 CLI	0.71	0.65	1.00	1.00
Gorenje gw 65 cl i	0.67	0.62	1.00	1.00
Electrolux EWS 1052 NDU	0.59	0.63	0.98	0.97
Electrolux EZB 53400 AX	0.35	0.58	0.98	0.97
Ariston LSTB 4B00 RU	0.62	0.60	1.00	0.99

REFERENCES

[1] Quentin André, Nicholas Reinholtz, and Bart De Langhe. "Can Consumers Learn Price Dispersion? Evidence for Dispersion Spillover Across Categories". In: *Journal of Consumer Research* 48.5 (2022), 756—774.

[2] Michael Baye and John Morgan. "Price Dispersion in the Lab and on the Internet: Theory and Evidence". In: *The RAND Journal of Economics* 35.3 (2004), p. 449. ISSN: 07416261. DOI: 10.2307/1593702. URL: http://doi.wiley.com/10.2307/1593702.

[3] Michael Baye and John Morgan. "Brand and price advertising in online markets". In: *Management Science* 55.7 (2009), pp. 1139–1151. DOI: 10.1287/mnsc.1090.1005.

[4] Michael R Baye, John Morgan, Patrick Scholten, et al. "Information, search, and price dispersion". In: vol. 1. Elsevier Amsterdam, 2006, pp. 323–375.

[5] Kathy Baylis and Jeffrey Perloff. "Price dispersion on the internet: Good firms and bad firms". In: *Review of Industrial Organization* 21.3 (2002), pp. 305–324. ISSN: 0889938X. DOI: 10.1023/A:1020415304514.

[6] Nikolay [dataset] Bazenkov, Marina Sandomirskaia, and Maria Kuznetsova. "Household applicances prices from a Russian online marketplace". In: *Mendeley Data* v2 (2019). DOI: 10.17632/wt93jy5f2x.2. URL: https://data.mendeley.com/datasets/wt93jy5f2x/2.

[7] Fernando Bernstein, Jing-Sheng Song, and Xiaona Zheng. ""Bricks-and-mortar" vs."clicks-and-mortar": An equilibrium analysis". In: *European Journal of Operational Research* 187.3 (2008), pp. 671–690. DOI: 10.1016/j.ejor.2006.04.047.

[8] Erik Brynjolfsson and Michael Smith. "Frictionless commerce? A comparison of Internet and conventional retailers". In: *Management Science* 46.4 (2000), pp. 563–585. ISSN: 00251909. DOI: 10.1287/mnsc.46.4.563.12061.

[9] Hsiao Chi Chen, James Friedman, and Jacques François Thisse. "Boundedly rational nash equilibrium: A probabilistic choice approach". In: *Games and Economic Behavior* 18.1 (1997), pp. 32–54. ISSN: 08998256. DOI: 10.1006/game.1997.0514.

[10] Yongmin Chen and Tianle Zhang. "Equilibrium price dispersion with heterogeneous searchers". In: *International Journal of Industrial Organization* 29.6 (2011), pp. 645–654. ISSN: 01677187. DOI: 10.1016/j.ijindorg.2011.03.007.

[11] Judith Chevalier and Austan Goolsbee. "Measuring prices and price competition online: Amazon. com and BarnesandNoble. com". In: *Quantitative marketing and Economics* 1.2 (2003), pp. 203–222.

[12] Data Insights, InSales, and PayU. *E-commerce in Russia 2014: annual report (in Russian)*. 2015. URL: http://www.datainsight.ru/files/DI_InSales_PayU-Ecommerce2014.pdf.

[13] Néstor Duch-Brown et al. "The impact of online sales on consumers and firms. Evidence from consumer electronics". In: *International Journal of Industrial Organization* 52 (2017), pp. 30–62.

[14] Glenn Ellison. *Bounded Rationality in Industrial Organization*. Tech. rep. MIT, 2006. URL: https://economics.mit.edu/files/904.

[15] Anindya Ghose and Yuliang Yao. "Using transaction prices to re-examine price dispersion in electronic markets". In: *Information Systems Research* 22.2 (2011), pp. 269–288.

[16] Yuriy Gorodnichenko, Viacheslav Sheremirov, and Oleksandr Talavera. "Price Setting in Online Markets: Does IT Click?" In: *Journal of the European Economic Association* 16.6 (2018), pp. 1764–1811. ISSN: 1542-4766. DOI: 10.1093/jeea/jvx050.

[17] Philip A. Haile, Ali Hortaćsu, and Grigory Kosenok. “On the empirical content of quantal response equilibrium”. In: *American Economic Review* 98.1 (2008), pp. 180–200. ISSN: 00028282. DOI: 10.1257/aer.98.1.180.

[18] Ernan Haruvy and Peter TL Popkowski Leszczyc. “Search and choice in online consumer auctions”. In: *Marketing Science* 29.6 (2010), pp. 1152–1164.

[19] Ernan Haruvy and Peter TL Popkowski Leszczyc. “Measuring the impact of price matching guarantees on price dispersion and willingness to pay in auctions”. In: *University of Alberta School of Business Research Paper* 2013-703 (2013).

[20] Han Hong and Matthew Shum. “Using price distributions to estimate search costs”. In: *The RAND Journal of Economics* 37.2 (2006), pp. 257–275.

[21] Greg Kaplan and Guido Menzio. “The morphology of price dispersion”. In: *International Economic Review* 56.4 (2015), pp. 1165–1206.

[22] Greg Kaplan et al. “Relative price dispersion: Evidence and theory”. In: *American Economic Journal: Microeconomics* 11.3 (2019), pp. 68–124.

[23] Charlie Lindgren et al. “Is intertemporal price discrimination the cause of price dispersion in markets with low search costs?” In: *Applied Economics Letters* 28.11 (2021), pp. 968–971.

[24] R. D. Luce. *Individual choice behavior: A theoretical analysis*. New York: Wiley, 1959.

[25] Daniel McFadden. “Econometric analysis of qualitative response models”. In: *Handbook of Econometrics*. Ed. by Z Griliches† and M D Intriligator. 1st ed. Vol. 2. Elsevier, 1984. Chap. 24, pp. 1395–1457. URL: https://econpapers.repec.org/RePEc:eee:ecochp:2-24.

[26] Richard D. McKelvey and Thomas R. Palfrey. “Quantal Response Equilibria for Normal Form Games”. In: *Games and Economic Behavior* 10.1 (1995), pp. 6 –38. ISSN: 0899-8256. DOI: https://doi.org/10.1006/game.1995.1023. URL: http://www.sciencedirect.com/science/article/pii/S0899825685710238.

[27] John Morgan, Henrik Orzen, and Martin Sefton. “An experimental study of price dispersion”. In: *Games and Economic Behavior* 54.1 (2006), pp. 134–158. ISSN: 08998256. DOI: 10.1016/j.geb.2004.07.005.

[28] Xing Pan, Brian T. Ratchford, and Venkatesh Shankar. “Price dispersion on the-internet: A review and directions for future research”. In: *Journal of Interactive Marketing* 18.4 (2004), pp. 116–135. ISSN: 10949968. DOI: 10.1002/dir.20019.

[29] Dieter Pennerstorfer et al. “Information and price dispersion: Theory and evidence”. In: *International Economic Review* 61.2 (2020), pp. 871–899.

[30] David Ronayne. “Price comparison websites”. In: *International Economic Review* 62.3 (2021), pp. 1081–1110.

[31] Steven Salop and Joseph Stiglitz. “Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion”. In: *The Review of Economic Studies* 44.3 (1977), p. 510. ISSN: 00346527. DOI: 10.2307/2296903. URL: https://www.jstor.org/stable/2296903.

[32] George J. Stigler. “The Economics of Information”. In: *Journal of Political Economy* 69.3 (1961), pp. 213–225. ISSN: 0022-3808. DOI: 10.1086/258464.

[33] Hal Ronald Varian. “A Model of Sales”. In: *American Economic Review* 70.4 (1980), pp. 651–59.

[34] Wenche Wang and Fan Li. “What determines online transaction price dispersion? Evidence from the largest online platform in China”. In: *Electronic Commerce Research and Applications* 42 (2020), p. 100968.

[35] Ruiliang Yan. "Pricing strategy for companies with mixed online and traditional retailing distribution markets". In: *Journal of Product & Brand Management* 17.1 (2008), pp. 48–56. DOI: https://doi.org/10.1108/10610420810856512.

[36] Yandex and Gfk. *Study of online-consumers in Russia, August – September 2014 (in Russian)*. Tech. rep. 2014. URL: https://cache-mskstoredata04.cdn.yandex.net/download.yandex.ru/company/figures/GfK_2014.pdf.

[37] Yandex and Gfk. *Study of online-consumers in Russia, September – November 2018 (in Russian)*. Tech. rep. 2018. URL: https://cache-mskm905.cdn.yandex.net/download.yandex.ru/company/figures/2018/gfk/yandex_market_gfk_2018.pdf.

