ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
"НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
"ВЫСШАЯ ШКОЛА ЭКОНОМИКИ"

НЕГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
"РОССИЙСКАЯ ЭКОНОМИЧЕСКАЯ ШКОЛА"(ИНСТИТУТ)

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Влияние трёхмесячной динамики цены на пользовательский спрос на рынке электронной коммерции

Программа Бакалавр экономики

Совместная программа по экономике НИУ ВШЭ и РЭШ

Автор:
А.С. НЕФФ

Научный руководитель:
Г.В. КОСЕНОК

Москва, 2021 г.

Аннотация

Пандемия коронавируса ускорила рост индустрии электронной коммерции и подчеркнула её важность для будущих покупателей. При этом до сих пор многое остаётся неизвестным относительно принципов, которыми руководствуются потребители при совершении покупок онлайн. Используя данные Яндекс Маркета, крупнейшего в России сайта для сравнения цен, мы исследуем влияние трёхмесячной динамики цены на мобильные телефоны на интерес пользователя к их покупке. Полученные результаты подтверждают влияние динамики цены на спрос, при этом падение цены после её роста обладает наиболее сильным эффектом на рост спроса среди рассмотренных нами паттернов эволюции цен. Эмпирически выделены две группы потребителей. Первая группа ограничена внутренним восприятием истории цен, в то время как вторая дополнительно получает информацию об исторической динамике цен с сайта. С использованием двух групп предложена эмпирическая стратегия разделения эффектов от внутренней и внешней информации о динамике цены. Установлено, что обладание внешней информацией о динамике цены снижает степень негативности реакции покупателя на последовательное падение цены и рост, следующий за падением. Качественное расхождение оцененных эффектов между двумя группами моделей показывает важность учёта действий покупателей, ушедших ни с чем. Это позволяет избежать смещения оценки параметров спроса из-за присутствия неслучайного отбора покупателей, основанного на их итоговом решении о покупке.

Abstract

The ongoing COVID-19 pandemic has accelerated the growth of the e-commerce industry and highlighted its importance for future customers. Meanwhile, a lot is still unknown about the principles of customer decision-making in the online environment. Using data from Yandex Market, the biggest Russian website for price comparison, we explore the impact of three-month price dynamics for smartphones on customer's interest in buying. The obtained results confirm that price dynamics do have an impact on demand, with price rise followed by its fall having the strongest effect on demand increase among studied patterns. Two groups of consumers are distinguished empirically. The first group only has an inner perception of price history, while the second group in addition gets precise information about historical price dynamics from the website. Employing these two groups, an empirical strategy is proposed that separates the effects from internal and external information about price dynamics. It was found that buyers who received external information about price dynamics react less negatively to persistent fall of price and price rise followed by fall. Qualitative discrepancies in estimated effects between two groups of models showed the importance of taking into account the actions of buyers who did not make a purchase. It helps us to avoid selection bias in estimates, resulting from the nonrandom exclusion of buyers based on their final purchase decision.

Contents

1 Introduction

The ongoing COVID-19 pandemic has given a significant boost to the already flourishing market of e-commerce. Customers all over the world were deprived of an opportunity to visit the shopping mall or restaurant during the long days of quarantine. Even those who have never shopped online before started using online marketplaces and food delivery services. Pandemic made sustainable changes in consumer behavior: in the US, 29% of surveyed people stated that they will never go back to offline shopping. Therefore, online shopping patterns become an important area to study both for academic and industrial researchers.

The story of e-commerce began in 1971 when the ARPANET packet-switching network was firstly used to arrange a cannabis[1] sale between students at Stanford and MIT. Starting from then, e-commerce was slowly evolving along with the development of the Internet[2] until the key years for contemporary e-commerce have come: in 1995, both Amazon and eBay were launched in the US, and in 1999 the Alibaba Group was established in China. Since then, the global B2C (business-to-customer) e-commerce market had grown from $150 billion back in 1999 to $4.3 trillion in 2020, accounting for as much as an astonishing 5.1% of global GDP.

From a global perspective, the Russian e-commerce market can be considered small yet extremely fast-growing. It had emerged in 1998 with the foundation of Ozon. During 2020, the Russian e-commerce market grew by 58% and constituted $37.4 billion, accounting for 2.5% of GDP[3], and is projected to triple up to $119.3 billion by 2024. For comparison, last year the global e-commerce market grew only by 28% and is projected to grow with 11% CAGR, compared to 34% CAGR projected for Russia. About a quarter of the Russian e-commerce market is now taken by universal online marketplaces. Marketplaces allow customers to order whatever they want choosing among offers from different vendors on one website. As 2020 showed, the market share of marketplaces in Russia is going to grow higher in the upcoming years. Becoming the main driver of e-commerce growth, Russian marketplaces, from which the biggest are Wildberries, Ozon, Aliexpress Russia, and Yandex Market, grew by 108% to $10.0 billion for the past year.

One of the cornerstones of any market is demand, and one conventional way to look at demand on the micro-level is to use the *reference price* concept[4]. Reference price is a price anticipated by a customer, which she compares with actual prices observed in the

[1] Cannabis is a popular recreational drug around the world. In the United States alone, it is believed that over 100 million Americans have tried cannabis.

[2] Internet started with ARPANET network in 1969 and remained fairly local until the development of World Wide Web (WWW) in 1989.

[3] The share of e-commerce in GDP in Russia is below the global one at the moment due to a later emergence and lower degree of internet penetration. In 2017, Russia took 52-th place in terms of internet availability, with about 75 percent of population having access to world web.

[4] The first important research on reference price was done by Winer (1986), and is discussed in details in Section 2.

shop. If the observed price appears to be lower than reference one, it increases the desire to buy. While the potential buyer becomes unhappy and wants to buy less if the observed price appears to be higher than the reference level. Among the important determinants of reference price, price history is the one that is frequently considered. History is usually included as an internal perception that the customer retrieves out of past experience.[5]

The understanding of price history's impact on demand is relevant for companies that sell their products online, as it may suggest the improvements for employed pricing strategies. Market researchers can also benefit from this study, as it may help them form more sophisticated models of consumer demand.

One of the first studies on reference price was done by Winer (1986), supporting empirically both reference price model for demand and price history as a valuable determinant of the reference price.[6] Putler (1992) showed the loss-aversion properties of reference price effects and impact reduction with the growth of reference price.[7] Based on the latter result, we suggested the usage of relative rather than absolute differences of actual price from reference in demand modeling. Briesch et al. (1997) performed a comprehensive empirical comparison of two stimulus-based and three memory-based models. They showed that the memory-based model with brand-specific past prices included is the best among all considered.

Among behavioral studies, the experiment[8] of Slonim and Garbarino (1999) also found the presence of prior reference price formation, which is followed by a comparison with the observed prices. Three experiments[9] set by Niedrich et al. (2001) showed that consumers compare the observed price of desired product against all of the prices in the set of comparable products rather than against the mean or range of available prices. These results highlight the complexity of customers' price perception, thus speaking in favor of elaborate demand modeling. Finally, Mazumdar et al. (2005) reviewed all the findings regarding reference price from behavioral and structural researches. They produced a classification for various properties of reference price formation, which we account for in our models.

E-commerce data allows us to differentiate between the customers' inner and external information about price dynamics.[10] Consequently, we can divide users into two groups.

[5] For example, the customer may remember the previous shopping experience for frequently purchased products. This inherent quality of price history accounting complicates the estimation of its effect on demand by making it highly heterogeneous and dependent on many unobservable factors that we cannot control for. For example, we know almost nothing about customer interactions with price information that happen between the shop visits.

[6] In Winer's study, the focus was placed on frequently purchased products, and price history took the form of the previous purchase price. For additional details see Section 2.

[7] Putler showed that the effect on buying probability of 1$ difference between the reference price and observed price diminishes with reference price growth.

[8] For details of experiment see Section 2.

[9] For details of experiments see Section 2.

[10] Inner price history information is the one customer has when she enters the shop. This information is approximate and can be obtained from friends, ads, previous purchases, and many other sources which

The first group only has an inner perception of price, while the second group in addition gets precise information about historical price dynamics on the website. Employing these two groups, we separate the effects from internal and external information about price dynamics. To infer the impact of price history we use two datasets built using the database of Yandex Market[11]. The aggregated dataset consists of successful clicks[12] that occurred during the second half of the year 2020. In turn, the view-based dataset has over 1.2 million customer views[13] of smartphone offers. The data on views is collected for 23 different smartphone models of 5 different brands and contains many characteristics of an offer, model, shop, and user session.

The study is focused on three-month patterns of price dynamics.[14] We develop a microeconomic static model of demand and treat data as cross-sectional since occurrences of the same user's multiple visits are rare.[15] Our main model of interest is Nested Logit, with other Logit models, such as Logit and Multinomial Logit, being used for preliminary analysis.[16]

We found that price history does affect the demand, with a price rise followed by its fall (*bell-shape drop*) being the most positive driver among studied patterns. Bell-shape drop in price increases the average user's click probability by 48% compared to a situation with a persistent fall in price. It was also shown that buyers who possess external information about the price dynamics react less negatively to persistent fall of price and price rise followed by fall. We also discuss the issues of working with data that contains only positive user interactions (clicks and purchases), as such data leads us to opposite estimated effects due to selection bias. The bias stems from the nonrandom exclusion of buyers from the sample based on their final purchase decision. We underline that for correct inference both positive and negative consumers' interactions should be present in data.

We highlight useful insights into consumer perception of different dynamic pricing patterns, along with the new possible way of incorporating the price history in the demand models. Therefore, we make a contribution to Consumer Economics and Marketing. In addition, our discussion regarding the appropriate selection of e-commerce data makes an implicit contribution to a vast group of empirical fields that can encounter such data. It is

we can not effectively control for. In contrast, external information is accurate and received by customers from the price comparison website. We are able to distinguish between inner and external information effects since our dataset captures user actions made on the website. In particular, we record interactions with the widget from the product card that shows a line plot with median prices for the last 6 months.

[11]Yandex Market, which was established in 2000, is the biggest Russian price comparison website.

[12]"Clicks" are navigations from a price comparison website to a partner's page. Partner is billed for all navigations, regardless of whether the purchase was made after all or not.

[13]"View" indicates such user interaction act, when she observes one of offers for a particular product.

[14]The month scale was chosen as the widget of interest shows the monthly dynamics. Three periods were chosen as a higher amount of periods would have created too many possible patterns that we are not able to account for in estimation due to multicollinearity issues in our data.

[15]For details see Section 5.

[16]Where Logits are not applicable due to data specifics, the OLS models are used.

also noteworthy, that the methodology suggested for three-month price history patterns can be easily re-used in other studies. The flexibility of methodology that allows one to incorporate any dynamic price patterns[17] into demand models makes this contribution fairly universal.

The exposition is organized as follows. Section 2 contains a brief review of theoretical and empirical works regarding the price history effects in the context of consumer demand estimation. In Section 3 one can find the description of price history patterns and datasets employed in estimation. The suggested empirical models are rigorously described in Section 4 and then the estimation strategy is shown in detail in Section 5. We continue in Section 6, presenting the results and their interpretation. Following the results, Section 7 contains robustness checks. After all, Section 8 summarizes the results of the undertaken research and discusses its limitations and prospects for future studies.

2 Literature Review

The research regarding customers' perception of the retail price had started back in the 1960s by Gabor and Granger (1961). Later, buyers' price expectations were discussed as "perceived price" (Monroe, 1973) and "price beliefs" (Erickson and Johansson, 1985). The *reference price* concept, which we actively use in our study, was introduced by Winer (1986). He defined reference price as a consumer's perceived current price of a brand, which one expects to observe at the point of purchase. Using the scanner data panel on coffee consumption of 1318 households[18], Winer established three important empirical results. Firstly, he demonstrated support for the reference price model in the case of frequently purchased products. Secondly, the deviation of observed price from the reference price had the suggested significant effects on purchases. Finally, price history in form of the price encountered during the previous purchase was shown to be an important source for present reference price formation.

Since 1986, extensive work on reference price had already been done both on the theoretical and empirical ground. Putler (1992) showed that customer response to gains and losses is likely to be heterogeneous and asymmetric.[19] These results thus suggested that reference price formation and utilization process is consistent with a concept of loss-aversion. He also showed that the effect on buying probability of 1$ difference between the reference price and observed price diminishes with reference price growth. Based on the above result, we suggest the usage of relative rather than absolute differences from

[17] One can include patterns with different scale or number of time periods and with different definition of drop, rise etc.

[18] All households were from a large midwestern city. Available information included purchasing histories of coffee on a daily basis, prices paid, demographic data, the store where coffee was bought, and daily advertising for coffee in local newspapers. The period of observation covered 429 days.

[19] Results were derived from the data which covered weekly per capita demand for eggs in Southern California, Jul 1981 – Jul 1983.

reference price in demand modeling.

Another important work was done by Briesch et al. (1997). They performed a comprehensive empirical comparison of two stimulus-based and three memory-based models of the reference price. Using data on four distinct popular products[20], authors showed the memory-based model with brand-specific past prices to be the best among all considered.

Among behavioral studies, the experiment of Slonim and Garbarino (1999) has also found evidence of buyer's prior reference price formation process followed by its comparison with prices observed in the shop. The authors engaged 445 students from business classes of two US universities. Each subject was given five consecutive simulated shopping decisions, in which students had to allocate a fixed budget between two ballpoint pens. The price of one brand remained constant at \$0.50, while the other brand's price was changing according to one of eight price orders. Both prices were ending at \$0.50 at the final decision. By asking how expensive the price-changing pen was at \$0.50, the authors were able to highlight the impact of so-called "perceived expensiveness". The results of this experiment suggest that the effect of reference price on demand passes through the product's perceived expensiveness. The higher is the observed price given the reference the more expensive a product appears to a customer, holding its characteristics constant. In turn, the demand for a pricey product is lower than for the one priced "fairly" in the eyes of a potential buyer.

Three experiments set by Niedrich et al. (2001) supported a proposition about sophisticated customer decision-making. They argued that consumers compare the observed price of the desired product against all of the prices in the set of comparable products rather than simply against the mean or range of available prices. These experiments involved students as well, showing them prices for airline tickets and carbonated beverages with discounts from 0% up to 60%. After seeing prices, students were asked to evaluate the price attractiveness on a nine-point scale. Prices were randomly drawn to be presented either in a sequential manner or in a simultaneous one. Using variation in presentation ways, the authors showed that "range effects" are stronger in the sequential presentation condition. In contrast, "frequency effects" turned to be stronger under simultaneous presentation.[21] Experimental results highlighted the complexity of customers' perception of price and thus spoke in favor of elaborate demand modeling.

A thorough review of all findings from both behavioral and structural pieces of research regarding reference price was done by Mazumdar et al. (2005). They produced a classification for various properties of the reference price formation process. Let us highlight a few of their conclusions that are particularly useful for our study.

[20]The data contained sales for peanut butter (4 brands), liquid detergent (5 brands), tissue (5 brands), and ground coffee (4 brands).

[21]*Range effects* manifested themselves in the impact of observed price's relative distance from the bounds of the set of all available prices. *Frequency effects* emerged as the impact of observed price's quantile in the set of all available prices.

Firstly, the price history is the strongest determinant of consumers' *internal reference price* (IRP)[22]. The more time has passed from the moment the price was encountered the weaker is the effect this price has on the IRP formation process. This fact justifies our choice of the shortened historical period for price patterns: three months instead of six. However, for durable products (as smartphones we employ in our study) the competitive price and attribute configurations also have a sizeable effect on IRP. This property demands an extensive set of contemporaneous prices and model characteristics to be included as covariates in our models to get a complete picture.

Secondly, when forming the reference price customers not only weigh past and present prices but are also able to infer a price trend. Through this channel, reference price has a significant effect on consumers' purchase-timing decisions. Due to these properties, we expect dynamic price patterns to have a significant effect on demand.

Authors also claim that consumers use both memory and external information while assigning weights to them. These weights depend on consumer and product characteristics. There are two factors suggesting that higher weight is placed on memory in the case of mobile phones. At first, many consumers have small consideration sets as they have their favorite brand with only a few models available within the desired price segment. Furthermore, the price level is generally high for smartphones, demanding the buyers to be thoughtful about observed prices. At the same time, the frequent promotions and high inter-purchase time speak in favor of higher weight being placed on external information[23]. That being said, we expect both memory and external information to play a significant role in the reference price formation process for smartphones.

Finally, the authors stated that the evidence for loss aversion regarding *sticker shock* (demand shock caused by a difference between the reference price and observed price) so far was mixed. Therefore, our work sheds additional light on this controversial issue.

Expanding the theoretical framework, Koszegi and Rabin (2006) allow the expected probability of buying to also affect the customer's reaction to sticker shock. An increase in the underlying likelihood of buying inflates consumer's disutility from product loss when the product is not purchased. Authors call this effect an "attachment effect" that strengthens customer's willingness to pay all things being equal. Along with the expected price, the likelihood of buying determines customers' shopping behavior. Unfortunately, we have not figured out the way of evaluating this perceived likelihood of a purchase from our data. However, we believe that not taking into account this characteristic does not distort our inference of the price history patterns' effect. That is likely to be true as the common reason for buying a new phone is the malfunction of an old one, which is likely

[22]The definition of *internal* reference price discussed by Mazumdar et al. (2005) coincides with our definition of reference price.

[23]External information consists of anything the customer may spot in the shop. Usually, it is the prices of desired product, its substitutes, and supplements. On the Yandex Market website, which we use in our study, customers can also employ the widget with information on the price history for the past 6 months.

exogenous relative to any price dynamics.

3 Data description

We use two datasets in this study. Among two, the *aggregated dataset* has more data at the expense of having only clicks – positive outcomes of user interaction with a particular offer. In our study, *click* indicates navigation from a price comparison website to a webpage of the shop that placed the offer. The shop is billed for all committed navigations, with an additional fee charged in case of a successful purchase later made by a user. The *view-based dataset* contains only interactions with a few smartphone models selected by us. But unlike the former dataset, it includes all customer interactions, both successful and not. *View* here identifies such user interaction act when she sees one particular offer out of all the offers available for a particular product. Due to plausible inherent selection bias in the aggregated dataset[24], the view-based dataset serves as our workhorse for the majority of estimated models.

Before we move to the description of datasets, we first need to classify the three-month patterns of price dynamics we employ in our study. We define three-month patterns using a triplet of most recent average monthly prices[25] relative to the moment of interaction. Let p_t be the product's average price in the month t, when the interaction (click for aggregated dataset or view for view-based one) happened. Correspondingly, p_{t-1} and p_{t-2} denote the average price one and two months before month t. Any three-month movement of price can thus be defined using two price changes: $\Delta_{t-1} = p_{t-1} - p_{t-2}$ and $\Delta_t = p_t - p_{t-1}$. Both changes can be either a rise ($\Delta_k > 0$), a fall ($\Delta_k < 0$) or a flat one ($\Delta_k = 0$). Hence there are 9 possible three-month price patterns, among which we distinguish the five, which are most commonly employed in our analysis: bell-shape drop, U-shape rise, persistent fall, persistent rise, and persistent flat. The complete classification of patterns is presented in Table 1.

In pattern classification, we consider the price deviation of any size, even if the change constitutes a small percent of the product's prior price. The rationale behind this division lies in principles of plot rendering implemented in the price history widget that we (and website users) utilize as the source of price history information. On this widget, the price axis consists only of prices between minimal and maximal average ones across the last 6 months. These rules of plot construction may result in a graph where even a small change is rendered as a substantial one. In data, we do not have enough variation in

[24]Issues regarding selection bias are discussed in Section 6.

[25]To construct the average monthly price we take the arithmetic average over median daily prices. The median price is a value such that at most half of the presented offers have a lower price than the proposed median and at most half have a higher price than the proposed median. Cited averaging algorithm replicates the way of price history formation on the widget employed in the study. The widget itself is discussed below in the current Section.

Table 1: The set of three-month patterns of price dynamics

Pattern	Δ_{t-1}	Δ_t
Persistent rise	> 0 (rise)	> 0 (rise)
Persistent fall	< 0 (fall)	< 0 (fall)
U-shape rise	< 0 (fall)	> 0 (rise)
Bell-shape drop	> 0 (rise)	< 0 (fall)
Persistent flat	$= 0$ (flat)	$= 0$ (flat)
Rise then flat	> 0 (rise)	$= 0$ (flat)
Fall then flat	< 0 (fall)	$= 0$ (flat)
Flat then rise	$= 0$ (flat)	> 0 (rise)
Flat then fall	$= 0$ (flat)	< 0 (fall)

price histories to make an additional distinction of patterns employing the visualization specifics. Therefore, we proceed with a more restrictive classification.

3.1 Aggregated dataset

The aggregated dataset contains the information about all clicks made by users from Central Federal District[26] via the desktop version of the Yandex Market website. The time period covers 6 months between 01.06.2020 and 30.11.2020. The total amount of clicks made by users during the studied period exceeds 66 million. In order to reduce the number of records, while keeping all the information about clicks, we group the observations, thus creating a seven-dimensional table. Each cell of the resulting table contains the number of clicks grouped according to values of the week, product category, product's average price bucket, offer percentage difference from average product's price, shop and model rating, along with an encountered pattern of price dynamics. From now on we call each cell a group, reflecting that it was constructed by grouping the observations. The number of clicks made within created groups becomes the outcome variable, with a larger amount of clicks manifesting higher customer demand. The descriptive statistics regarding grouping variables and amount of groups along each dimension are shown and discussed below in this subsection.

Table 2 and Table 3 show dominance of price buckets above 3000 RUB and expensive product categories, such as electronics and home appliance. The distribution of click shares is fairly uniform, indicating that we constructed reasonable price buckets.

Regarding patterns of price dynamics, Table 4 shows that slightly more clicks are made when the price is persistently rising or falling. Many clicks are also made with bell-shape

[26]Central Federal District is the most populated federal district in Russia with more than 39 million people living in 18 provinces. It is entirely located in the European part of Russia, occupying about 19% of it and about 4% of the whole Russian territory. The biggest cities in Central Federal District are Moscow, Voronezh and Yaroslavl.

Table 2: Aggregated dataset – average price bucket

Price bucket	Share of all clicks	Number of groups	Avg. clicks per group
0–1000 RUB	18.6%	678,141	18.1
1000–3000 RUB	18.0%	545,790	21.7
3000–10000 RUB	20.5%	487,872	33.6
10000–25000 RUB	18.0%	314,125	43.1
25000+ RUB	24.9%	229,994	51.8
Total: 5	100%	2,255,922	29.3

Note: Share of all clicks is computed as the total amount of clicks in all groups with a given fixed price bucket divided by the total amount of clicks in the sample. The number of groups represents the total amount of groups that contain the corresponding price bucket in one of the dimensions. Average clicks per group measure is calculated as the total amount of clicks in all groups with a given fixed price bucket divided by a number of such groups.

Table 3: Aggregated dataset – product category

Product category	Share of all clicks	Number of groups	Avg. clicks per group
Electronics (incl. Smartphones)	22.2%	203,710	72.1
Home Appliance	15.5%	188,131	54.3
Computer equipment	13.3%	187,189	47.1
Construction and repair	11.1%	271,366	26.9
Automobile	5.7%	146,396	25.6
Household	5.5%	211,259	17.1
Health	5.4%	119,999	29.6
Children	5.3%	178,115	19.7
Beauty	4.3%	173,557	16.2
Sport and recreation	3.8%	134,241	18.7
Country house and garden	2.9%	108,407	17.6
Pet	1.9%	85,432	14.3
Grocery	1.2%	86,144	9.4
Clothing, shoes and accessories	1.1%	85,853	8.8
Leisure and entertainment	0.6%	48,269	8.5
Appliance	0.3%	28,854	6.0
Total: 16	100%	2,255,922	29.3

Note: Share of all clicks is computed as the total amount of clicks in all groups with a given fixed product category divided by the total amount of clicks in the sample. The number of groups represents the total amount of groups that contain the corresponding product category in one of the dimensions. Average clicks per group measure is calculated as the total amount of clicks in all groups with a given fixed product category divided by a number of such groups.

and U-shape price patterns, which are expected to happen around periods of sales.[27]

Table 4: Aggregated dataset – three-month price dynamics

Price dynamics pattern	Share of all clicks	Number of groups	Avg. clicks per group
Persistent rise	26.3%	439,142	39.5
Persistent fall	22.9%	414,570	36.5
U-shape rise	20.2%	410,852	32.4
Bell-shape drop	19.4%	407,807	31.3
Persistent flat	5.6%	180,936	20.5
Rise then flat	2.2%	130,034	11.0
Fall then flat	1.6%	106,186	10.2
Flat then rise	0.9%	81,673	7.2
Flat then fall	0.9%	84,722	7.1
Total: 9	100%	2,255,922	29.3

Note: Share of all clicks is computed as the total amount of clicks in all groups with a given fixed dynamics pattern divided by total amount of clicks in the sample. The number of groups represents the total amount of groups that contain the corresponding dynamics pattern in one of the dimensions. Average clicks per group measure is calculated as the total amount of clicks in all groups with a given fixed dynamics pattern divided by a number of such groups.

Figure 1 demonstrates the weekly distribution of clicks over time. The patterns observed in the sample are consistent with expected demand behavior in retail: it is usually higher in the second (Apr-Jun) and fourth (Oct-Dec) quarters of a year.[28]

The descriptive statistics regarding the remaining groupping characteristics (offer percentage difference from average product's price, shop rating and model rating) can be found in Tables A.2 – A.4 situated in Appendix 9.1.

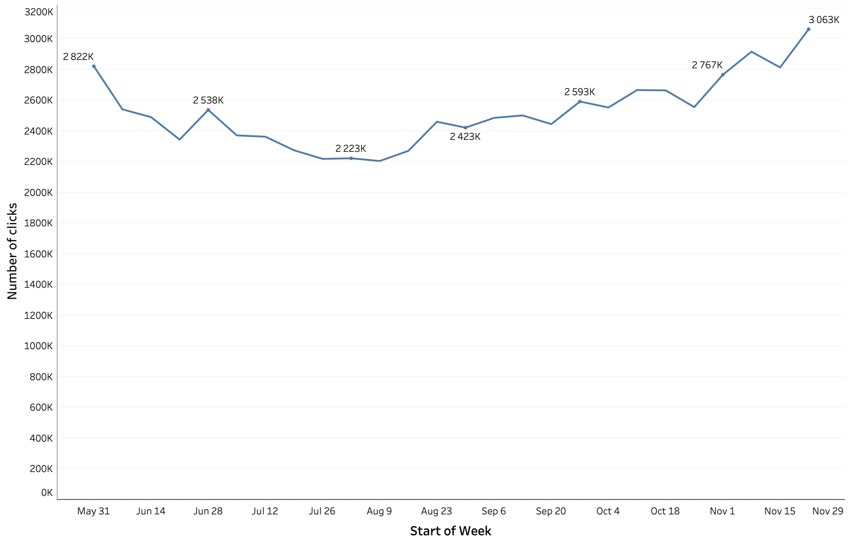
At the end of aggregated dataset's creation process, we got rid of groups with no clicks made.[29] We did this since most of the groups without clicks demonstrate not the customer's disappointment regarding a particular combination of 7 grouping characteristics but the supply absence of corresponding offers. If we let such zero-groups stay, then

[27] We expect a bell-shape drop to appear during the sale itself and a U-shape rise to appear in the next period when prices rise to former levels.

[28] The boost of demand in the fourth quarter is driven by customers buying gifts for Christmas and New Year and stirred by huge November sales such as Black Friday. In the following first quarter, customers restrain themselves from active shopping, while having holidays at the beginning of January and starting to work later to compensate for all the money spent during New Year's Eve. Then, in the second quarter, the may holiday come into place, with people being ready to spend money again. Analogously, the second quarter is followed by the low season in the third quarter, when most people are either on vacation or working to pay for the vacation.

[29] Out of potential 1.2 billion groups we have only 0.1% left – 2.3 million groups. The big part of occurred eradication attributes to a large number of groups for offer percentage difference from product's average price. We decided both to not trim observations with high percentage differences and to not group them into one atom. The first measure would result in the loss of information about existing outliers, which are not completely irrelevant since some customers did click on them. The second measure would result in the overestimation of effect from offer price difference since the atom would have a higher amount of clicks than his neighbors with a more favorable price.

Figure 1: Click distribution within time period (01.06.2020 – 30.11.2020)



Note: The total number of weeks in the sample is 27. The descriptive statistics in the alternative form can be found in Table A.1 in Appendix 9.1.

our estimates would have a negative bias regarding unpopular characteristics. However, despite we get rid of the bias mentioned above, aggregated dataset nevertheless contains the selection bias which is discussed in detail in Section 5.

3.2 View-based dataset

View-based dataset, which is restricted in product variety, user population, and time period, consists of 1,250,000 views of offers for 23 popular smartphone models made by users from Moscow via the Yandex Market desktop website. The considered time period covers around 2.5 months, from 17.09.2020 to 30.11.2020.[30] Within one observation, each view is linked with characteristics of an offer, model, shop, and user session. Each observation also contains the results of user interaction with an offer, represented by the following two dummies. The first dummy indicates whether the click for this particular offer occurred. The second dummy indicates the presence of purchase committed on the partner's website after a click. While we allow the click to happen without the following purchase, the opposite is prohibited by the design of our variables.

[30] To preserve variation in price histories, we chose the widest available period, for which all the characteristics of our interest are available to be linked to a view. Acquired variation is large enough to estimate models of our interest but does not allow us to include location-level differentiation in effects. So, in order to lower the noise associated with regional heterogeneity of customers, we proceed with data gathered on users from one particular region: Moscow. As the capital and the most populated city in Russia (with over 12 million people), Moscow has the highest variety of products and shops, and generates the most views and clicks across all Russian provinces. Smartphones were chosen as they are the signature products for Yandex Market, and thus allow us to maximize the richness of data, holding constant the period and region. We keep only the most popular 23 models as popular models are subjected to the highest amount of shops' competition. It allows popular models to experience decent variation in prices and thus in price patterns even given our relatively small time frame.

Table 5 contains summary statistics for important dummies.[31] It highlights an important difference between offline and online shopping. The conversion from view to navigation in the online environment is quite low (about 2%) and conversion from view to actual purchase is microscopic (about 0.07%). Largely due to this property the estimates made using purchases as the target variable are too noisy to detect a significant effect. Therefore, in our main estimations, we use clicks as the proxy for customer demand, while inspecting in Section 7 whether a click is good as a proxy.

Table 5: Descriptive statistics – View-based dataset, dummy variables

	Sum	Mean
User made a click	25,170	0.0202
User made a purchase	888	0.000714
User has seen the widget with price history	154,451	0.124
Offer from Top-6 block	634,674	0.510
Offer from Default block	328,739	0.264
Brand 1 (Apple)	353,942	0.284
Brand 2 (Samsung)	294,534	0.237
Brand 3 (Xiaomi)	361,533	0.291
Brand 4 (Honor)	103,359	0.083
Brand 5 (Huawei)	130,823	0.105
Number of views	1,244,191	

Note: The sum of dummy variables indicates the number of successful outcomes. Mean is the sample probability of encountering the record with a corresponding characteristic among all views. As the dataset contains only 5 listed brands, the mean of all brands adds up to 1 and the sum of brand successful observations adds up to a total number of observations.

The set of chosen smartphones has a significant variation in prices, from inexpensive Xiaomi models to costly Apple and Samsung fresh models. Table 6 shows that average prices for models included in our sample vary from 10490 RUB (about $135) to 82990 RUB (about $1,075).

All nine potential price history patterns occur in data with positive probability (Table 7). Unfortunately, the view-based dataset contains visits made during only 2.5 months. Therefore, in order to preserve some variation for the rarely encountered mixed flat patterns[32] we pool them together with the persistent rise dynamics. In that way, we isolate four patterns that we are particularly interested in studying and comparing: U-shape rise,

[31] *Click* is navigation from a price comparison website to a shop page of the partner who placed the offer. *Purchase* indicates the order made on the partner's website after the click. The price history widget is discussed above in Section 3. The offer seen by a customer may be located in one of 3 blocks on the product card. *Premium block* (1 offer) is placed on the page's top, right next to product photos, and contains the offer from the partner who won the auction and paid to be placed there. *Default block* (1 offer) marked as "low price" is located right under premium one and contains the best offer according to a ranking algorithm. *Top-6* block is located below photos and main characteristics and consists of 6 best offers according to the ranking algorithm (in the absence of the offer placed as default one).

[32] Patterns with flat price in one period and shift in another appeared to be the rarest (Table 7).

Table 6: Descriptive statistics – View-based dataset, prices

	Obs.	Mean	SD	Min	Max
Offer price, rub		36,151	21,113	7,215	526,590
Average model price, rub		37,581	21,171	10,490	82,990
Minimal model price, rub		32,660	19,105	7,215	82,349
Premium offer price, rub		37,497	21,644	8,763	108,861
Default offer price, rub		32,910	19,353	1,097	89,990
Number of views	1,244,191				

Note: Minimal and average model prices are calculated among all available offers for a particular model. Premium offer is placed on the page's top, right next to product photos, and is an offer from the partner who won the auction to be placed there. As expected, the minimal model price is slightly lower than all other prices. The average offer price is lower than the mean of average model prices as the price comparison service promotes offers with lower prices, all other things being equal. Default offer, as an offer that was considered the most attractive by ranking algorithm, usually has a low price slightly above the minimal one. Premium offer is slightly more expensive than an average one for a given model because the promotions are often used to distinguish a particular offer from others, in an attempt to create an additional value for the customer, which can then be extracted via setting a relatively higher price.

bell-shape drop, persistent fall, and persistent flat. Four selected patterns cover 75% of all views. However, when the appearance of a pattern is intersected with the user's interaction with the price history widget, the frequency of each pattern occurrence appears to be quite low.

As in the aggregated dataset, the persistent fall pattern is encountered more often than U-shape rise and bell-shape drop. At the same time, in the view-based dataset persistent rise pattern is only a third in views[33] relative to being the most clicked in the aggregated dataset.[34]

In Figure 2 the reader can observe the various price history patterns as they appear in the sample. The detailed description of variables contained in our datasets can be found in Appendix (9.2 for the aggregated dataset and 9.3 for the view-based one).

4 Model

To allow for some degree of interchangeability among different mobile phone models we use the Nested Logit framework as the control point for modelling.[35] Each view of an offer is supplemented with characteristics of three following dimensions: i – shop, j –

[33]The most popular pattern in the view-based dataset is persistent fall and the second most popular is U-shape rise.

[34]Less frequent appearance of persistent rise dynamics can be explained by the specificity of the smartphone market, which is characterized by active product obsolescence. The fast development of new models makes it harder for shops to increase the prices for old devices as the new ones are always coming to the market.

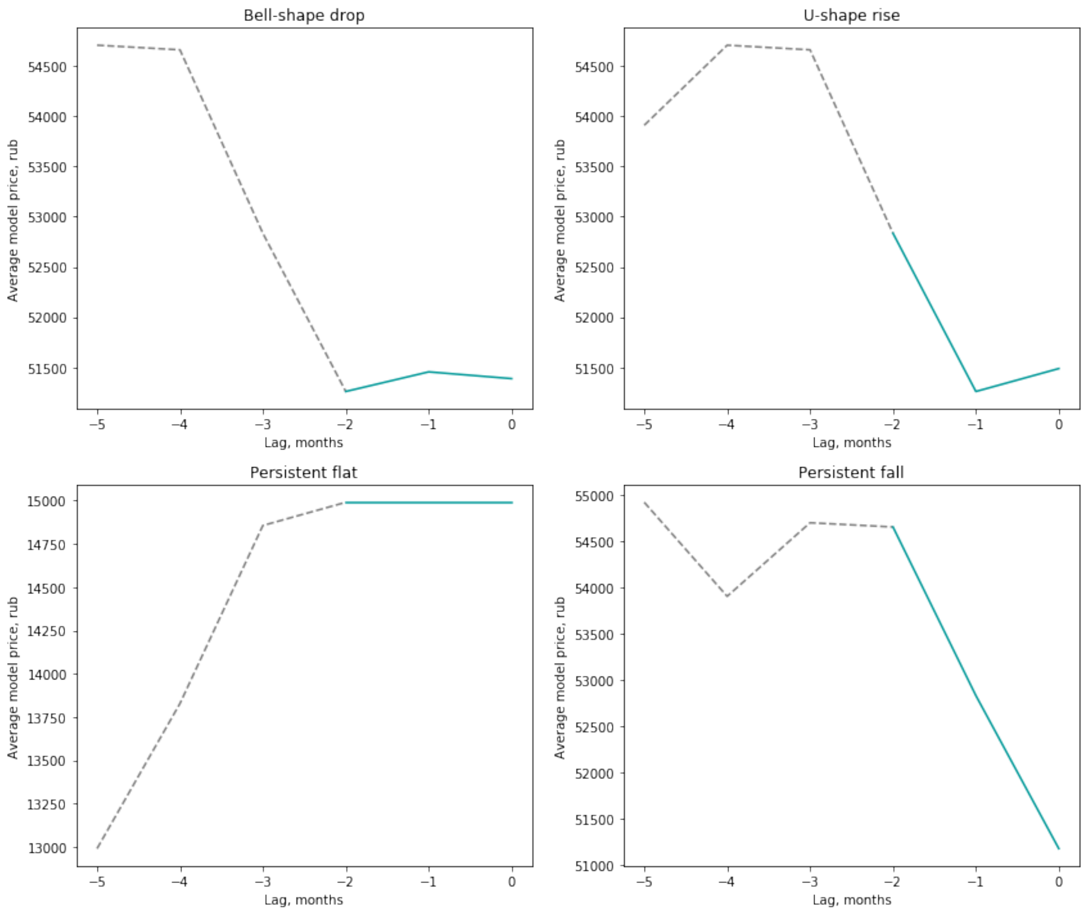
[35]For an extensive description of Nested Logit Model see pages 48-49 and 362-364 from Aguirregabiria (2019).

Table 7: Descriptive statistics – View-based dataset, price history dummies

	Sum	Mean
Persistent fall	370,112	0.297
U-shape rise	342,301	0.275
Bell-shape drop	206,661	0.166
Persistent flat	23,566	0.019
Persistent fall (interaction)	47,026	0.038
U-shape rise (interaction)	42,704	0.034
Bell-shape drop (interaction)	24,665	0.020
Persistent flat (interaction)	2,998	0.002
Number of views	1,244,191	

Note: The sum of dummy variables indicates the total number of pattern occurrences in the sample. Mean is the sample probability of encountering the observations with a corresponding pattern among all views.

Figure 2: Examples of price history patterns



product, and h – user. In total, there are 1484 shops ($i = 1, ..., 1484$), 23 smartphones ($j = 1, ..., 23$), and 339772 visits ($h = 1, ..., 339772$) in the view-based dataset.[36]. The total amount of shop-product pairs (i, j) with an encountered offer is 7122 out of 37648 possible combinations. It means that during visits included in our sample an average shop has shown offers for at least 4 smartphones among selected ones.

Utility function .

We assume that each customer makes the buying decision using the utility function which depends on the comparison of actual observed price (shop specific, as product card has offers from various shops) with the expected (reference) one. We also allow the utility to depend on both model and shop characteristics.

$$U_{ijh} = \alpha_1 p_{ijh} + \alpha_2(p_{ijh} - RP_{jh}) + \alpha_3 DSH_{jh} + \beta_1 x_i^T + \beta_2 x_j^T + \beta_3 x_h^T + \varepsilon_{ijh} \tag{1}$$

Among factors affecting customer's utility, p_{ijh} is offer price (shop specific) that the user encountered, RP_{jh} is reference price (which includes price history patterns as one of determinants), DSH_{jh} is a dummy indicating whether user saw the price history widget, x_i – vector with two shop specific characteristics (stars and count of reviews), x_j – vector with six model specific characteristics (represented as overall user rating of a model as well as ratings for certain aspects of smartphone), x_h – vector (in our case degenerate) with one user session characteristic (session length). Regarding the error term, $\varepsilon_{ijh} = \lambda\varepsilon_{igh}^{(1)} + \varepsilon_{ijh}^{(2)}$, where g indicates the group of model (Apple-branded, Samsung-branded or Chinese smartphones) and $\varepsilon_{igh}^{(1)}$ and $\varepsilon_{ijh}^{(2)}$ are i.i.d. Extreme Value type 1 variables.[37]

DSH_{jh} is a dummy of particular importance for us, as it allows us to classify customers into two groups.[38] The first group ($DSH_{jh} = 0$) only has the inner perception of price, while the second ($DSH_{jh} = 1$) in addition gets precise information about historical price dynamics from the price history widget located on a model card. This distinction allows us to split the reference price into two components (internal and external)[39] and estimate the partial effects.

Reference price formation .

Reference price RP_{jh} formation process is assumed to be the same for all customers

[36]We treat data as cross-sectional one, i.e. consider all visits made by distinct users, since occurrences of the same user's multiple visits are scarce. For details see Section 5.

[37]Error term contains various unobserved characteristics of shop, model, user, and her session on the website. λ is a parameter, which measures the strength of correlation between different groups.

[38]DSH_{ij} is represented in data as variable *HasSeenAvgPriceProxyFlag* from the view-based dataset. For detailed description regarding this variable see Appendix 9.3.

[39]Internal component of reference price includes various information the user has when she enters the shop. In contrast, the external component contains the information the user receives from the price comparison website, including the price history information from the corresponding widget.

(conditioned on the fact of observing the price history widget) for a given model

$$RP_{jh} = \gamma_1 \cdot AP_{jh} + \gamma_2 \cdot POP_{jh} + w_{14} \cdot DOP_{jh} + w_{15} \cdot MinP_{jh} + \\ + \omega_1 \cdot DPat_{jh}^{(1)} + \omega_2 \cdot DPat_{jh}^{(2)} + \omega_3 \cdot DPat_{jh}^{(3)} + \omega_4 \cdot DPat_{jh}^{(4)} + \\ + \omega_5 \cdot DPat_{jh}^{(1)} \cdot DSH_{jh} + \omega_6 \cdot DPat_{jh}^{(2)} \cdot DSH_{jh} + \omega_7 \cdot DPat_{jh}^{(3)} \cdot DSH_{jh} + \omega_8 \cdot DPat_{jh}^{(4)} \cdot DSH_{jh} \tag{2}$$

Let us remind that j denotes the smartphone model and h is an index for agent.[40] The first four variables from the right-hand side, which are described in Table 6, have the following meaning. AP_{jh} stands for current median price, POP_{jh} is a price of premium offer, DOP_{jh} is a price of default offer for the model, and $MinP_{jh}$ is a minimum price for the model. $DPat_{jh}^{(k)}$ (where $k \in \{1,2,3,4\}$ is a row number from Table 7) are dummies for the encountered pattern of price history. The interactions with DSH_{jh} are included to allow for heterogeneity of the reference price formation process for users with and without explicit information about price history.

Decision-making process .

Let $A = \{(i, j) : i = 1,...,1484;\ j = 1,..., 23\} \cup (0, 0)$ be the set of all shop-product pairs, with pair $(0, 0)$ corresponding to an outside option with zero utility. In turn, let A_{av} be the set of all pairs (m, n) from A for which the shop m has an available offer for model n. Following the Logit framework, the customer's decision regarding purchase is done using the following log-odds:

$$\mathbb{P}\{j_h = j,\, i_h = i\} = \frac{\exp(U_{ijh})}{1 + \sum_{m,n} \mathbf{1}\{(m,n) \in A_{av}\} \exp(U_{mnh})} = V_{ijh} + \varepsilon_{ijh} \tag{3}$$

where V_{ijh} is defined from formula (1).

5 Estimation

In this study, we have two questions to answer. Firstly, we test whether three-month price dynamics have an impact on customer demand. Secondly, we check whether this effect (of price dynamics on demand) is different for customers with and without external information[41] about price history.

In what follows, we treat data in a cross-sectional manner since occurrences of the

[40]Despite we assumed users' similarity in terms of reference price formation process, they enter the website at different times. Consequently, users encounter distinct prices and price history patterns. For each view, we record all related variables exactly at the moment of view occurrence, without any extrapolation.

[41]We define external information as one, which is received directly from the price history widget on the website.

same user's multiple visits during the selected time period are rare.[42] To perform a comprehensive test of proposed hypotheses we estimate 6 different models and compare their results. The first two models are designed for OLS estimation and employ the aggregated dataset, while the following four models are designed for Logit estimation and use the view-based dataset. Throughout the estimation, we use clicks to proxy customer demand.[43] The discussion concerning the same six models estimated using actual purchases as the target variable can be found in Section 7, which contains a set of robustness checks for our results.

The first two models are estimated using OLS and the aggregated dataset. Stemming from the aggregated structure of data, the number of clicks in a group serves as the dependent variable. The set of available covariates lacks the variables related to the price history widget, as we don't have information about user sessions in the aggregated dataset. The estimated regression is the following for **Model 1** (OLS without controls)[44]

$$\begin{aligned}Clicks = \alpha_0 TotalClicks + \sum_{n=1}^{4} \alpha_n PriceBucketDummy^{(n)} + \\ + \beta_1 DPat^{(1)} + \beta_2 DPat^{(2)} + \beta_3 DPat^{(3)} + \beta_4 DPat^{(4)} + \varepsilon\end{aligned}$$

and the following for **Model 2** (OLS with controls):

$$\begin{aligned}Clicks = \alpha_0 TotalClicks + \sum_{n=1}^{4} \alpha_n PriceBucketDummy^{(n)} + \\ + \alpha_5 ODAvg + \alpha_6 ShopRating + \sum_{k=1}^{15} \gamma_k CatDummy^{(k)} + \\ + \beta_1 DPat^{(1)} + \beta_2 DPat^{(2)} + \beta_3 DPat^{(3)} + \beta_4 DPat^{(4)} + \varepsilon\end{aligned}$$

Among the variables that are not described above, $TotalClicks$ is a total amount of clicks among all groups inside the particular week of click, $PriceBucketDummy^{(n)}$ is a dummy indicating whether the average price belongs to particular one of 5 price buckets (here, n stands for a row number from Table 2), $ODAvg$ stands for offer percentage difference from the model average price, $ShopRating$ is shop's rating on a five-point scale (from 0 to 5, with 1 point step), and $CatDummy^{(k)}$ is a dummy indicating whether the product belongs to one of 16 product categories[45] (here, k stands for a row number from

[42] Among users from our view-based dataset, 73% had exactly one session and 87% had two sessions or less.

[43] For details regarding the definition of click and the decision to use clicks as demand proxy see Section 3.

[44] Standard errors are clustered at the level of a week of click to account for possible shifts in demand that affect the whole website. Such shifts can be caused by sales or vice versa by the onset of the low season.

[45] Each product belongs to exactly one product category among distinguished.

Table 3).

Moving to the view-based dataset, which has a binary dependent variable (dummy indicating whether there was a click on a particular viewed offer), we switch to Logit specification. At first, we estimate two Logit models, one without controls and another with controls included. The regression, which is estimated using maximum likelihood criteria, is the following for **Model 3** (Logit without controls)[46]

$$\begin{gathered}
ClickDummy_{ijh} = \alpha_1 OfferPrice_{ijh} + \alpha_2 DSH_{jh} + \\
+ \beta_1 DPat^{(1)}_{jh} + \beta_2 DPat^{(2)}_{jh} + \beta_3 DPat^{(3)}_{jh} + \beta_4 DPat^{(4)}_{jh} + \\
+ \omega_1 DPat^{(1)}_{jh} \cdot DSH_{jh} + \omega_2 DPat^{(2)}_{jh} \cdot DSH_{jh} + \omega_3 DPat^{(3)}_{jh} \cdot DSH_{jh} + \omega_4 DPat^{(4)}_{jh} \cdot DSH_{jh} + \varepsilon_{ijh}
\end{gathered}$$

and the following for **Model 4** (Logit with controls):

$$\begin{gathered}
ClickDummy_{ijh} = \alpha_1 OfferPrice_{ijh} + \alpha_2 DSH_{jh} + \\
+ \beta_1 DPat^{(1)}_{jh} + \beta_2 DPat^{(2)}_{jh} + \beta_3 DPat^{(3)}_{jh} + \beta_4 DPat^{(4)}_{jh} + \\
+ \omega_1 DPat^{(1)}_{jh} \cdot DSH_{jh} + \omega_2 DPat^{(2)}_{jh} \cdot DSH_{jh} + \omega_3 DPat^{(3)}_{jh} \cdot DSH_{jh} + \omega_4 DPat^{(4)}_{jh} \cdot DSH_{jh} + \\
+ \alpha_3 PriceDifFromAvg_{ijh} + \alpha_4 PriceDifFromMin_{ijh} + \\
+ \alpha_5 PriceDifFromDefault_{ijh} + \alpha_6 PriceDifFromPremium_{ijh} + \\
+ \alpha_7 ModelRating_{jh} + \alpha_8 ModelReviewsCount_{jh} + \sum_{k=1}^{5} \gamma_k ModelFeatureRating^{(k)}_{jh} + \\
+ \alpha_9 ShopRating_{ih} + \alpha_{10} ShopGradesCount_{ih} + \\
+ \alpha_{11} SessionLength_{ijh} + \sum_{n=1}^{4} \psi_n BrandDummy^{(n)}_{j} + \varepsilon_{ijh}
\end{gathered}$$

Among the variables, which are not described above, four $PriceDifFromX_{ijh}$ variables denote offer percentage difference from corresponding model-level prices (average, minimal, and prices of offers from default and premium blocks for the model respectively), $ModelRating_{jh}$ (from 1 to 5, with 0.5 point step) and $ModelReviewsCount_{jh}$ are similar to their shop analogs, $ModelFeatureRating^{(k)}_{jh}$ is a five-point scale rating (from 1 to 5, with 0.1 point step) of certain model certain features[47], $ShopGradesCount_{ih}$ is a number of reviews left by customers regarding a particular shop (a proxy for shop's popularity), $SessionLength_{ijh}$ is a number of seconds passed from the start of the session until the offer view, and $BrandDummy^{(n)}_{j}$ are dummies which indicate that model belongs to certain brand listed in Table 5.

Thereafter, we come down to Multinomial logit, which allows the user to have only

[46]Standard errors in all 4 Logit-based models are clustered at the level of visit id, which incorporates all views made by a user during one session.

[47]Duration of autonomous usage, memory volume, camera quality, overall performance, and screen quality.

one purchase among 5 selected brands. For estimation purposes, the outcome variable is turned into a categorical one, which takes the number of a chosen brand, with 0 value corresponding to an outside option – going away without acquisition of a new smartphone. In case of successful click occurrence for more than one brand during the same session, we duplicate all views the user made during this visit. For instance, if during the same visit models from both brands m and n are clicked, we create two similar sets of views. The first set contains outcome click on brand m and the second one contains outcome click on brand n. Due to low variation, we had to omit *ShopRating* from the set of controls used in models 3 and 4 in order to tackle multicollinearity issues.[48] Therefore, **Model 5** is represented by the following equation:

$$\begin{gathered}
ClickBrandDummy_{ijh} = \alpha_1 OfferPrice_{ijh} + \alpha_2 DSH_{jh} + \\
+ \beta_1 DPat_{jh}^{(1)} + \beta_2 DPat_{jh}^{(2)} + \beta_3 DPat_{jh}^{(3)} + \beta_4 DPat_{jh}^{(4)} + \\
+ \omega_1 DPat_{jh}^{(1)} \cdot DSH_{jh} + \omega_2 DPat_{jh}^{(2)} \cdot DSH_{jh} + \omega_3 DPat_{jh}^{(3)} \cdot DSH_{jh} + \omega_4 DPat_{jh}^{(4)} \cdot DSH_{jh} + \\
+ \alpha_3 PriceDifFromAvg_{ijh} + \alpha_4 PriceDifFromMin_{ijh} + \\
+ \alpha_5 PriceDifFromDefault_{ijh} + \alpha_6 PriceDifFromPremium_{ijh} + \\
+ \alpha_8 ModelReviewsCount_{jh} + \sum_{k=1}^{5} \gamma_k ModelFeatureRating_{jh}^{(k)} + \\
+ \alpha_9 ShopRating_{ih} + \alpha_{10} ShopGradesCount_{ih} + \alpha_{11} SessionLength_{ijh} + \varepsilon_{ijh}
\end{gathered}$$

The estimation ends with our least restrictive model – Nested Logit. The nested structure allows customers to firstly choose which group of smartphones they are interested in (Apple-branded, Samsung-branded, or one of the Chinese well-known brands – Xiaomi, Honor, and Huawei) and then pick a model inside the chosen group that matches them the best. All offers for one model (from various shops), which the user has seen during the session, are gathered up in one record with averaged characteristics of price, rating, and so on. The dataset was then modified in such a way that each session contains the outside option represented by a model with zero characteristics. Unfortunately, due to low variation, we had to omit almost all controls from regression in order to tackle

[48] All coefficients are generally consistent between models 3, 4, and 5, which indicates that *ShopRating* is unlikely to cause an omitted variable bias for the coefficients of interest. Another explanation regarding the absence of effect from offer-specific variables on product-specific coefficients is discussed in Section 6.

multicollinearity issues.[49] Therefore, **Model 6** is represented by the following equation:

$$
\begin{aligned}
ClickDummy_{jh} = \alpha_1 OfferPrice_{jh} + DSH_{jh} + \\
+ \beta_1 DPat_{jh}^{(1)} + \beta_2 DPat_{jh}^{(2)} + \beta_3 DPat_{jh}^{(3)} + \beta_4 DPat_{jh}^{(4)} + \\
+ \omega_1 DPat_{jh}^{(1)} \cdot DSH_{jh} + \omega_2 DPat_{jh}^{(2)} \cdot DSH_{jh} + \omega_3 DPat_{jh}^{(3)} \cdot DSH_{jh} + \omega_4 DPat_{jh}^{(4)} \cdot DSH_{jh} + \\
+ \alpha_{11} SessionLength_{jh} + \varepsilon_{jh}
\end{aligned}
$$

6 Results

Tables 8 and 9 present the main estimation results (the estimated parameters for control variables from Models 3-6 are presented in Appendixes 9.4 – 9.6). We compare four price history patterns that are explicitly included in regression against five omitted ones. The major omitted pattern is a persistent rise since views of other patterns (flat in one period and non-flat in another) together constitute only about 18% of all views with one of omitted price history patterns encountered.

Models that are estimated using the aggregated dataset (Table 8) suggest that a history of persistent fall draws the highest customer demand. U-shape rise is the second-best pattern for customers, and a bell-shape drop is the third preferable option among selected patterns. All three best selected patterns have positive effects on demand relative to omitted ones. Only persistent flat price history is estimated to have a negative impact on demand. We think this coefficient is estimated to be negative because the products with exact flat price history are barely encountered by customers,[50] thus generating fewer clicks on average due to supply shortage.

Regarding Models 3-6 designed for Logit estimation (Table 9), we observe that all four selected price history patterns have significant effects on the probability of making a click. For a median user, who does not look at the price history widget, the bell-shape drop in price has a positive effect on demand. At the same time, the other three selected patterns have a negative effect compared to omitted ones. The second-best pattern among selected is U-shape rise, the third-best is persistent fall and the worst pattern is again persistent flat.

The relative magnitude and signs of price history effects are radically different between the models estimated using the aggregated dataset and the view-based one. Therefore, it is important to discuss the selection bias that, in our opinion, spoils the estimations that are employing the aggregated dataset. As only successful clicks appear in the aggregated dataset, the more product offers do exist with specific price history, the more clicks these offers get on average. That being said, the important variable *Views* with a significant

[49] The results of this omission are discussed in Section 6.

[50] For smartphones from the view-based dataset we observe this situation explicitly in Table 7.

Table 8: Estimation results, aggregated dataset

VARIABLES	(1) OLS *Clicks*	(2) OLS with controls *Clicks*
BellShapeDrop	6.916***	7.207***
	(0.930)	(0.955)
UShapeRise	7.862***	8.152***
	(0.965)	(1.041)
PersistentFlat	−4.961***	−5.265***
	(0.859)	(0.877)
PersistentFall	12.29***	11.97***
	(1.692)	(1.748)
TotalClicks	8.12e-06***	8.90e-06***
	(4.08e-07)	(3.20e-07)
Price_1000-3000_Dummy	3.68***	1.63***
	(0.300)	(0.294)
Price_3000-10000_Dummy	15.69***	13.03***
	(0.536)	(0.471)
Price_10000-25000_Dummy	25.22***	20.46***
	(0.921)	(0.802)
Price_25000+_Dummy	34.17***	25.65***
	(0.763)	(0.589)
OfferDiffFromAvgPrice (ODAvg)		−0.0822**
		(0.030)
ShopRating		9.49***
		(0.166)
Constant	−8.813***	−16.55***
	(1.281)	(1.284)
Observations	2,255,922	2,255,922
Product category dummies	no	yes
R2	0.005	0.015

*Note: The outcome variable is the number of clicks made in a particular group. Among dummies for price buckets, the omitted one is [Price 0-1000 Dummy]. All employed variables are rigorously described in Section 5. Clustered (by a week of click) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Table 9: Estimation results, view-based dataset

VARIABLES	(3) Logit *ClickDummy*	(4) Logit with controls *ClickDummy*	(5) Multinomial Logit[1] *ClickBrandDummy*	(6) Nested Logit[3] *ClickDummy*
OfferPrice	$-1.75e\text{-}06^{***}$	$-3.69e\text{-}06^{***}$	-0.000160^{***}	-0.000329^{***}
	(3.71e-07)	(1.69e-06)	(8.07e-07)	(2.00e-06)
DSH	0.0457	0.170^{***}	0.600^{***}	0.312^{***}
	(0.0433)	(0.0442)	(0.0798)	(0.0412)
BellShapeDrop	0.174^{***}	0.0857^{***}	0.212^{***}	0.139^{***}
	(0.0234)	(0.0261)	(0.0569)	(0.0264)
BellShapeDrop $\times$	-0.0537	-0.0812	-0.0382	-0.0320
	(0.0697)	(0.0710)	(0.118)	(0.0650)
UShapeRise	-0.0611^{***}	-0.197^{***}	0.0328	-0.172^{***}
	(0.0218)	(0.0248)	(0.0455)	(0.0244)
UShapeRise $\times$	0.0542	0.0666	0.0282	0.181^{***}
	(0.0614)	(0.0624)	(0.0977)	(0.0580)
PersistentFlat	-0.510^{***}	-0.869^{***}	0.563	
	(0.0779)	(0.0100)	(0.426)	
PersistentFlat $\times$	-0.0718	-0.0741	1.397^{***}	
	(0.257)	(0.258)	(0.447)	
PersistentFall	0.00872	-0.149^{***}	0.0414	-0.261^{***}
	(0.0219)	(0.0283)	(0.0522)	(0.0303)
PersistentFall $\times$	0.157^{***}	0.143^{**}	0.232^{**}	0.177^{***}
	(0.0607)	(0.0615)	(0.100)	(0.0560)
Constant	-3.840^{***}	22.82^{***}	-272.1^{***}	...
	(0.0180)	(2.320)	(10.569)	...
Observations	1,244,191	1,244,191	1,245,546[2]	5,873,606[4]
Controls	no	yes	yes	yes
R2	0.0012	0.0511	0.2776	0.0608

Note: The outcome variable is a dummy indicating whether the click on viewed offer happened. For Multinomial Logit in case of click, dummy takes the value of the clicked brand index. All variables and are rigorously described in Section 5. Mark × denotes the interaction with DSH dummy. Clustered (by user session) standard errors in parentheses.

**** p<0.01, ** p<0.05, * p<0.1.*

[1] *Results are presented only for one selected group: Xiaomi-branded phones. For full results see Table A.6 in Appendix 9.5.*

[2] *The outcome variable was tracked and extrapolated through the same session. In case of successfull clicks for multiple brands during one session we duplicated the session.*

[3] *Dummy for flat price history was omitted due to multicollinearity. For full results see Table A.7 in Appendix 9.6.*

[4] *Dataset was transformed in such a way that it fits the nested logit structure: each visit has exactly one alternative with a successful click, either for one of the products or for an outside option. Other available alternatives are represented each in separate observation with outcome variable equal to 0.*

positive effect on the dependent variable *Clicks* is inherently omitted in the first two models. Consequently, the estimates presented in Table 8 are likely to overstate the true effect for frequently encountered price history patterns and understate for rare ones. The mentioned bias can not be resolved without switching to more detailed data containing information regarding views, which is provided in our view-based dataset.

In Model 4, almost all controls (Table A.5 in Appendix 9.4) have expected signs. Differences of offer price relative to various references[51] have the strongest negative effect among all variables, which agrees with theoretical predictions discussed in Section 2.[52] At the same time, the effect from the offer absolute price itself becomes insignificant when controlled for relative differences, which supports fragmentation of users into groups done via product nesting in Model 6.[53] Model quality (rating) and popularity (measured as a number of reviews left by users) affect the demand positively. However, the fact that shop rating negatively affects the demand (all else being equal) suggests that there exist shop-specific omitted variables (for example, cost and speed of delivery that the shop offers). Despite these omitted shop-specific variables may affect the estimated coefficients for included controls, they are unlikely to cause bias in product-specific coefficients of interest.[54]

Moreover, the significant coefficients of interest are consistent in signs among all four models. It speaks in favor of the reliability of our results, suggesting that our qualitative inference made with Models 3-6 is robust both to relaxing some of the restrictive assumptions made by ordinary Logit models and to the partial omission of covariates. Therefore, we consider the view-based dataset and so the results of Models 3-6 more robust ones and focus on them in further discussion about estimated price history effects. Nevertheless, we think that due to an ambiguity about the fulfillment of the CIA condition[55] the magnitude of effects needs to be validated in further research.

Interpreting the results of Model 6, we can say that the price history of bell-shape drop increases the z-score of click probability by 0.139. Consequently, we expect the bell-shape drop to increase the click probability for median customer, who did not see the price history widget, by 0.3 percentage points.[56] This increase seems small at first, but so far

[51]Average price, minimal price, prices of default and premium offers serve as captured references.

[52]In particular, the significant role devoted to contemporaneous sources for referencing was predicted by Mazumdar et al. (2005).

[53]Inside her price segment, the customer cares not about the price itself (as the phone's characteristics justify the cost) but about an opportunity to buy the same phone cheaper.

[54]It is true if we assume that shop does not use price history when setting the offer conditions such as terms of delivery. We believe in this assumption since terms of delivery are usually constrained only by the shop's infrastructure and thus are the same among different smartphones sold by the same shop.

[55]CIA stands for Conditional Independent Assumption and is crucial for regression to have a causal interpretation of results. While dealing with encountered multicollinearity issues in estimation we had to omit lots of covariates that we believed to be important for customer decision-making. As there exist plausible channels for correlation between omitted covariates and price history, we can not reject the hypothesis that CIA is not satisfied in our case.

[56]Since the average probability of click equals to 0.020, z-score can be calculated from log odds as

we compared the bell-shape drop with a mixture of omitted price histories. If we compare bell-shape drop with persistent fall history, we get the 0.7 pp increase, which increases the click probability by 48%, from 0.016 to 0.023. Therefore, we consider the estimated effect to be large in the economic sense. It is also noteworthy that our point estimates are also consistent with the loss aversion property suggested by Putler (1992). The increase in z-score caused by bell-shape drop price history (0.139) is lower in absolute terms than the decrease caused by the inverse pattern, U-shape rise (-0.172).

The positive effect of the bell-shape drop is consistent in sign and significant for all Models 3-6 and for two specific brands in Multinomial Logit (Xiaomi and Huawei). The effect of the inverse pattern, U-shape rise, is negative, while also being consistent in sign and significant for Models 3, 4, and 6.[57] Despite having a negative effect on demand in general, the persistent fall pattern appears to have a positive effect on demand for phones made by Huawei. We have a similar situation with the persistent flat pattern, which has a positive effect on demand for Samsung-branded phones. Unfortunately, the persistent flat price pattern was the least frequent among four selected patterns and therefore had to be omitted in Nested Logit due to multicollinearity issues. Thus the best available estimate for the effect of the persistent flat price pattern comes from Model 4.[58] Brand-specific differences in estimated effects (Table A.6) suggest some degree of consumers' heterogeneity that is mostly left beyond the scope of this study.

One valid explanation why bell-shape drop leads to the strongest demand increase is the customer's feeling of gain, even though no explicit discount is made in this case. Following the mean-reversion logic, the customer may expect that in a short time the price will probably return to "normal" high levels. This expectation stimulates her not to postpone the purchase. The same mean-reversion logic can be applied to the U-shape rise in price, which makes the customer perceive the current price to be higher than the "normal" one. However, for the U-shape rise, the opposite effect may also exist. The increase in price may also stimulate to not postpone buying if the price is expected to go even higher following the upward trend set by the preceding rise. Our results suggest that the mean-reverting logic dominates among customers.

The estimated difference in effect that is related to observing the price history widget is insignificant for bell-shape drop.[59] At the same time, the difference is significant for

$\ln(\frac{0.020}{1-0.020}) = -3.892$. With 0.139 increase, z-score goes up from -3.892 to -3.753. Inverting the calculation, we get the new probability of 0.023.

[57] In Multinomial Logit the effect is significant for three specific brands: Samsung, Honor, and Huawei.

[58] As we said that all models are fairly consistent in sign predictions among themselves, we can rely on the estimated sign for persistent flat pattern, even though it is inferred not from Model 6.

[59] This insignificance may be explained in two different ways. One explanation is based on assumption that people retrieve the price history information effectively enough using their memory only. The second explanation assumes the resulted insignificance comes from the imprecision of our estimates. The interaction of two rare events (seeing price history widget and observing one of the patterns of interest) occurs not frequently enough to allow us to capture the effect precisely. We encountered the same issue of result interpretation duality with estimated effects on purchases, as there are only 888 purchases out

U-shape rise and persistent fall patterns. Despite having a negative effect on average, U-shape rise and persistent fall patterns have significantly positive coefficients on interaction. In the case of the U-shape rise, the interaction effect even makes the point estimate of overall effect positive for users who saw the price history widget. It serves as important evidence in favor of our expectation that people receive the additional inputs for reference price formation from the widget. One possible explanation of the positive interaction effect is the after-sale behavior.[60] The average customer, who already spent her money during the sale, is less willing to buy a new phone until she finds for sure that the situation is especially conducive to this.[61]

7 Robustness checks

To check whether clicks serve as a good proxy for customer demand we re-estimated Models 1-5 using actual purchases as the outcome variable.[62] The estimation results are presented in Tables A.8 – A.11 in Appendixes 9.7 – 9.9. Estimated effects match well with the effects we got using clicks but are less precise due to the rare occasion of purchases in data.[63] Since the results look similar for clicks and purchases, the use of clicks in our main specifications is justified in order to get more precise estimates for effects of interest.

We also verify that difference between results we got using aggregated dataset and view-based dataset can not be attributed to differences in rules we used to fill our samples. In particular, datasets differ in the time period (six months in the aggregated dataset instead of three in view-based one) and contained products (all products vs. smartphones). We re-estimate the first two models on the aggregated dataset while restricting it for the same time period and products as the view-based dataset has. The inconsistency of signs and relative coefficient magnitude between Models 1-2 and Models 3-6 persists.[64] It

of 1,250,000 views captured in the sample.

[60]U-shape rise is likely to occur after the sale since if the sale was in the previous period, then the past price would almost surely be lower than two periods ago, setting the base for U-shape pattern in case of after-sale price revert and for persistent fall in case of sale continuation.

[61]The stimulus in this setup may come from the price increase if the customer expects the price to grow even higher in future due to price mean-reverting process after sale. However, the stimulus may vice versa happen if a customer does not see an increase and thus decides to buy before the potential price increase due to inflation or currency instability. The first example explains why the coefficient on interaction is positive for U-shape rise, while the second example shows that the similar coefficient can be positive for persistent fall.

[62]Price comparison website tracks the purchases that occurred at the partner's website after the click and bill them additionally. Unfortunately, we were unable to identify Model 6 for purchases due to a lack of variation in outcomes.

[63]The problem can potentially be solved in further research either by including more observations or (vice versa) by restricting the observations. The latter can be done by excluding the observations that are unlikely to be conscious enough. For example, one can exclude observations without meaningful interactions with an interface as they can be just miss-clicks. We decided to not exclude any observations in the current study to not miss the important information about outliers that do exist in the general population.

[64]The results and code for replication are available upon request sent to aneff@nes.ru.

suggests that our reasoning behind the importance of selection bias is valid and justifies the decision to use the view-based dataset as our workhorse.

Finally, we re-estimate our main models (Models 3-6 for click dummy as outcome variable) using the view-based dataset with trimming made on session length and offer price percentage difference from average product's price. That being said, we get rid of views with unusual session length or offer price difference belonging to the top and bottom 1% percentiles. By doing that, we check that our results are not driven by certain marginal offers or users. The magnitude and significance of estimated coefficients persist, thus ensuring the robustness of our results to the presence of outliers.[65]

8 Conclusion

We found the evidence of significant effect the three-month price history has on customer demand. For example, bell-shape drop (the best pattern among selected ones), increases the probability of click by 48% compared to persistent fall of price. It is also the only selected pattern that has a positive effect relative to a combination of five omitted patterns. The second-best pattern among selected ones is U-shape rise, the third-best is persistent fall and the worst pattern is persistent flat.

Moreover, for some patterns, we captured the mitigating effect that the external price history information has on the total effect from price history. In the case of U-shape rise and persistent fall, observing the price dynamics explicitly (employing price history widget) leads to a less negative reaction, which even turns to be slightly positive in total for U-shape rise. In our estimates, we use clicks as a proxy for demand, while showing that estimations done using actual purchases lead to similar but less precise results.

We also highlight the issues regarding data collection for researches related to e-commerce. Models that are estimated using the aggregated dataset, which contains only positive user interactions (clicks), led us to different conclusions due to persistent selection bias. This bias stems from occurred nonrandom exclusion of buyers based on their final buying decision. We indicate that for correct inference both positive and negative interactions of customers should be present in a sample. Consequently, we use the view-based dataset for correct inference about price history effects.

We conclude that our causal inference is robust both to relaxing some of the restrictive assumptions made by ordinary Logit models and to the partial omission of covariates. Nevertheless, we think that due to ambiguity about the fulfillment of the CIA condition the estimated magnitude of effects needs to be validated in further research. There are a few issues left beyond the scope of this study that can be addressed. Firstly, it is useful to differentiate between different degrees of price changes. For the sake of preserving variation by keeping the number of different patterns small, we treated the drop of aver-

[65]The results and code for replication are available upon request sent to aneff@nes.ru.

age price by 10 and 10000 roubles the same. Meanwhile, one can let the classification of patterns depend on the slope of price change. In our study, the decision on keeping the changes binary is justified by plot rendering principles used for the price history widget.[66] The second issue of potential interest is a seasonal change of effects' magnitudes, as users may behave differently during sales compared to regular shopping. Unfortunately, our data do not have enough time variation to explore the seasonality properly. Finally, our data lacks decent information on demographics. At the same time, the demographical differences in price history sensitivity and other issues regarding customer heterogeneity are important for the implementation of our results into optimal pricing strategies.[67] Therefore, we suggest the usage of Mixed Logit in further research to explore the heterogeneity of price history effects properly.

[66]For details, see Section 3.

[67]Distinct products have a different target audience, some of which may have the effect that differs fundamentally from the average one we estimated in this study.

References

Aguirregabiria, V. (2019). Empirical industrial organization: Models, methods, and applications.

Briesch, R. A., L. Krishnamurthi, T. Mazumdar, and S. P. Raj (1997). A Comparative Analysis of Reference Price Models. *Journal of Consumer Research 24*(2), 202–214.

Erickson, G. M. and J. K. Johansson (1985). The Role of Price in Multi-Attribute Product Evaluations. *Journal of Consumer Research 12*(2), 195–199.

Gabor, A. and C. W. J. Granger (1961). On the Price Consciousness of Consumers. *Journal of the Royal Statistical Society 10*(3), 170–188.

Koszegi, B. and M. Rabin (2006). A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics 121*(4), 1133–1165.

Mazumdar, T., S. Raj, and I. Sinha (2005). Reference price research: Review and propositions. *Journal of Marketing 69*(4), 84–102.

Monroe, K. B. (1973). Buyers' subjective perceptions of price. *Journal of Marketing Research 10*(1), 70–80.

Niedrich, R. W., S. Sharma, and D. H. Wedell (2001). Reference Price and Price Perceptions: A Comparison of Alternative Models. *Journal of Consumer Research 28*(3), 339–354.

Putler, D. S. (1992). Incorporating reference price effects into a theory of consumer choice. *Marketing Science 11*(3), 287–309.

Slonim, R. and E. Garbarino (1999). The effect of price history on demand as mediated by perceived price expensiveness. *Journal of Business Research 45*(1), 1–14.

Winer, R. S. (1986). A Reference Price Model of Brand Choice for Frequently Purchased Products. *Journal of Consumer Research 13*(2), 250–256.

9 Appendix

9.1 Additional descriptive statistics: aggregared dataset

Table A.1: Aggregated dataset – week of click

Week of click	Share of all clicks	Number of groups	Avg. clicks per group
2020-06-01	4.3	84,356	33.5
2020-06-08	3.8	80,221	31.7
2020-06-15	3.8	80,570	30.9
2020-06-22	3.6	78,407	29.9
2020-06-29	3.8	86,089	29.5
2020-07-06	3.6	81,085	29.3
2020-07-13	3.6	82,661	28.6
2020-07-20	3.4	81,709	27.8
2020-07-27	3.4	84,294	26.3
2020-08-03	3.4	81,491	27.3
2020-08-10	3.3	82,161	26.8
2020-08-17	3.4	82,929	27.4
2020-08-24	3.7	84,993	29.0
2020-08-31	3.7	87,909	27.6
2020-09-07	3.8	86,110	28.9
2020-09-14	3.8	86,522	28.9
2020-09-21	3.7	86,741	28.2
2020-09-28	3.9	92,572	28.0
2020-10-05	3.9	86,254	29.6
2020-10-12	4.0	86,031	31.0
2020-10-19	4.0	86,781	30.7
2020-10-26	3.9	89,383	28.6
2020-11-02	4.2	85,314	32.4
2020-11-09	4.4	87,891	33.2
2020-11-16	4.3	88,742	31.7
2020-11-23	4.6	90,727	33.8
2020-11-30	0.7	43,979	10.0
Total: 27	100%	2,255,922	29.3

Note: Share of all clicks is computed as the total amount of clicks in all groups with a given fixed week of click divided by total amount of clicks in the sample. The number of groups represents the total amount of groups that contain the corresponding week of click in one of the dimensions. Average clicks per group measure is calculated as the total amount of clicks in all groups with a given fixed week of click divided by a number of such groups.

Table A.2: Aggregated dataset – offer percentage difference from average product's price

Offer percentage difference	Share of all clicks	Number of groups	Avg. clicks per group
$[-100\%, -17.5\%)$	17.3%	585,800	19.5
$[-17.5\%, -12.5\%)$	13.8%	142,371	64.1
$[-12.5\%, -7.5\%)$	14.6%	150,969	63.9
$[-7.5\%, -2.5\%)$	12.8%	153,950	55.1
$[-2.5\%, 2.5\%)$	21.2%	210,782	66.5
$[2.5\%, 7.5\%)$	4.2%	123,124	22.3
$[7.5\%, 12.5\%)$	2.0%	99,141	13.4
$[12.5\%, 17.5\%)$	1.1%	81,215	9.2
$[17.5\%, +\infty]$	2.3%	510,857	3.0
Total: 1141	100%	2,255,922	29.3

Note: Share of all clicks is computed as the total amount of clicks in all groups with a given fixed offer percentage difference divided by total amount of clicks in the sample. The number of groups represents the total amount of groups that contain the corresponding offer percentage difference in one of the dimensions. Average clicks per group measure is calculated as the total amount of clicks in all groups with a given fixed offer percentage difference divided by a number of such groups. Users do not observe the offer percentage difference explicitly but we assume that they approximate it using the information available to them on the product page: the price of offer they look at and the average product price. Except for marginal meta-groups (denoted as $[-100\%, -17.5\%)$ and $[17.5\%, +\infty]$ in Table), which represent sets of groups, all groups are presented in the way they appear in data, i.e. with 5 percentage point increments in the difference.

Table A.3: Aggregated dataset – shop rating

Shop rating	Share of all clicks	Number of groups	Avg. clicks per group
0	0.2%	61,554	2.3
1	0.002%	832	1.8
2	0.1%	40,982	2.6
3	0.8%	120,425	3.7
4	32.4%	849,511	25.0
5	66.5%	1,182,618	37.3
Total: 6	100%	2,255,922	29.3

Note: Share of all clicks is computed as the total amount of clicks in all groups with a given fixed shop rating divided by total amount of clicks in the sample. The number of groups represents the total amount of groups that contain the corresponding shop rating in one of the dimensions. Average clicks per group measure is calculated as the total amount of clicks in all groups with a given fixed shop rating divided by a number of such groups. Shop rating is constructed as the average of users' reviews with a grade on a 5-point scale (from 1 to 5). Zero stars are assigned to shops with not enough reviews made for the website to be confident in rating values. The most popular shops have 4-5 stars as expected due to the endogeneity of customer choice: only good shops can grow popular.

Table A.4: Aggregated dataset – model rating

Model rating	Share of all clicks	Number of groups	Avg. clicks per group
1.0	0.01%	1,843	1.8
1.5	0.01%	4,795	2.1
2.0	0.04%	12,242	2.1
2.5	0.2%	42,901	2.5
3.0	0.5%	88,483	3.6
3.5	1.9%	199,391	6.4
4.0	8.2%	344,555	16.0
4.5	55.3%	579,957	66.0
5.0	16.2%	453,806	25.9
Total: 9	100%	2,255,922	29.3

Note: Share of all clicks is computed as the total amount of clicks in all groups with a given fixed model rating divided by total amount of clicks in the sample. The number of groups represents the total amount of groups that contain the corresponding model rating in one of the dimensions. Average clicks per group measure is calculated as the total amount of clicks in all groups with a given fixed model rating divided by a number of such groups. Model rating is constructed as the average of users' reviews with a grade on a 5-point scale (from 1 to 5). The most popular models have a rating between 4 and 5 as expected due to endogeneity of customer choice: the better is the model the more customers are lured by its high rating.

9.2 Detailed description of aggregated dataset

The aggregated dataset contains the information about all clicks made by users from the Central Federal District via the desktop version of the Yandex Market website. The time period covers 6 months between 01.06.2020 and 30.11.2020. Each cell of the resulting table contains the number of clicks grouped according to values of the week, product category, average price bucket, offer percentage difference from average product's price, shop and model rating, along with an encountered pattern of price dynamics. The total amount of observations (groups) is equal to 2,255,922. The total amount of clicks made within all groups is 66,020,849 and the total amount of purchases equals 3,059,727. The description of the initial (raw) dataset's variables along with variables that emerged during the data transformation process is presented below:

Clicks – total amount of clicks in particular group

Purchases – total amount of purchases in particular group

ClickWeek – date of click occurence (attributed by the start of the corresponding week)

ClickCategoryName – product category of the clicked offer's product. For estimation purposes we split this variable into the dummies for each category, which results in 16 variables, from *Cat1Dummy* to *Cat16Dummy*.

AvgPriceBucket – bucket of average price of chosen product (one of 5). For estimation purposes we split this variable into the following 5 dummies, one for each bucket:

Price_0-1000_Dummy, *Price_1000-3000_Dummy*, *Price_3000-10000_Dummy*, *Price_10000-25000_Dummy*, *Price_25000+_Dummy*
OfferDiffFromAvgPrice – percentage difference of clicked offer from the product's average price
TotalClicks – total amount of clicks among all groups inside the particular *ClickWeek*
TotalPurchases – total amount of clicks among all groups inside the particular *ClickWeek*
ShopRating – rating of clicked offer's shop on five-point scale (0 to 5, steps by 1)
ModelRating – rating of clicked offer's product on five-point scale (1 to 5, steps by 0.5)

For definitions of price history patterns see Section 3.
BellShapeDrop – dummy: product had the bell-shape drop price history at the moment of click
UShapeRise – dummy: product had the U-shape rise price history at the moment of click
PersistentFlat – dummy: product had the persistent flat price history at the moment of click
PersistentFall – dummy: product had the persistent fall price history at the moment of click
FallThenFlat – dummy: product had the fall then flat price history at the moment of click
RiseThenFlat – dummy: product had the rise then flat price history at the moment of click
FlatThenFall – dummy: product had the flat then fall price history at the moment of click
FlatThenRise – dummy: product had the flat then rise price history at the moment of click

9.3 Detailed description of view-based dataset

The view-based dataset consists of 1,250,000 views of offers for 23 popular smartphone models made by users from Moscow via the Yandex Market desktop website. The considered time period covers around 2.5 months, from 17.09.2020 to 30.11.2020. Within one observation, each view is linked with characteristics of an offer, model, shop, and user session. The record also contains the results of interaction, represented by the following two dummies. One dummy indicates whether the click for this particular offer occurred and another dummy indicates the presence of purchase. Total amount of observations (views) is 1,244,191, with 25,170 clicks and 888 purchases. The description of the initial (raw) dataset's variables along with variables that emerged during the data transformation process is presented below:
EventDate – the date when the view happened.

EventTime – the datetime when the view happened.

UserID – the hashed identificator of user.

VisitID – the hashed identificator of user session (visit).

ClickDummy – dummy: the viewed offer was clicked later.

PurchaseDummy – dummy: the viewed offer was purchased later.

ClickBrandDummy – the value i denotes that the viewed offer of brand i was clicked later (with 0 corresponding to an outside option).

PurchaseBrandDummy – the value i denotes that the viewed offer of brand i was purchased later (with 0 corresponding to an outside option).

OfferPrice – price of viewed offer.

OfferRawPosition – offer position according to ranking algorithm; Offer with position 0 is the best offer and is usually placed in the default offer block.

OfferTypeID – offer position type: 1 - block “Top 6”, 2 - default block, 3 - premium block.

ShopID – the hashed identificator of shop.

HasSeenAvgPriceFlag – dummy: user has seen the price history widget. Imprecise measure, which is based on the fact that the widget was rendered.

HasSeenAvgPriceProxyFlag (DSH) – dummy: user has seen the price history widget. More precise measure (which we use in our study), which is based on the fact that content below the widget was rendered.

ShopRating – rating of viewed offer’s shop on five-point scale (0 to 5, steps by 1).

ShopGradesCount – number of ratings left by users for viewed offer’s shop (proxy for shop prominence).

PriceDifFromAvg – percentage difference of viewed offer from the product’s contemporaneous average price *ModelAvgPrice*.

PriceDifFromMin – percentage difference of viewed offer from the product’s contemporaneous minimal price *ModelMinPrice*.

PriceDifFromDefault – percentage difference of viewed offer from the price of contemporaneous default offer for the product *ModelDefaultOfferPrice*.

PriceDifFromPremium – percentage difference of viewed offer from the price of contemporaneous premium offer for the product *ModelPremiumOfferPrice*.

SessionLength – number of seconds passed from start of the session until the offer view.

ModelID – hashed identificator of smartphone model

ModelBrandID – encoded product brand: 1 - Apple, 2 - Samsung, 3 - Xiaomi, 4 - Honor, 5 - Huawei.

ModelAvgPrice – product contemporaneous average price (among all available offers) at the moment of view

ModelAvgPriceLag1 – product average price in the previous month $(t-1)$ relative to the moment of view (month t)

ModelAvgPriceLag2 – product average price in month $t-2$ relative to the moment of view (month t)

ModelAvgPriceLag3 – product average price in month $t-3$ relative to the moment of view (month t)

ModelAvgPriceLag4 – product average price in month $t-4$ relative to the moment of view (month t)

ModelAvgPriceLag5 – product average price in month $t-5$ relative to the moment of view (month t)

ModelRating – rating of viewed product's shop on five-point scale (1 to 5, steps by 0.5).

ModelReviewsCount – number of reviews left by users for viewed offer's product (proxy for product prominence).

ModelCountOffers – contemporaneous number of available offers for product at the moment of view.

ModelFeature1AutDur – product's rating for duration of autonomous usage (rated by users, 1 to 5, steps by 0.1).

ModelFeature2MemVol – product's rating for memory volume (rated by users, 1 to 5, steps by 0.1).

ModelFeature3Camera – product's rating for camera quality (rated by users, 1 to 5, steps by 0.1).

ModelFeature4Performance – product's rating for overall performance (rated by users, 1 to 5, steps by 0.1).

ModelFeature5Screen – product's rating for screen quality (rated by users, 1 to 5, steps by 0.1).

For definitions of price history patterns see Section 3.

BellShapeDrop – dummy: product had the bell-shape drop price history at the moment of click

BellShapeDropX – dummy: product had the bell-shape drop price history at the moment of click AND user had observed this on the price history widget

UShapeRise – dummy: product had the U-shape rise price history at the moment of click

UShapeRiseX – dummy: product had the U-shape rise price history at the moment of click AND user had observed this on the price history widget

PersistentFlat – dummy: product had the persistent flat price history at the moment of click

PersistentFlatX – dummy: product had the persistent flat price history at the moment of click AND user had observed this on the price history widget

PersistentFall – dummy: product had the persistent fall price history at the moment of click

PersistentFallX – dummy: product had the persistent fall price history at the moment of click AND user had observed this on the price history widget

9.4 Estimation Results: Models 3-4 for clicks.

Table A.5: Estimation results (for clicks) for control variables, Models 3-4

VARIABLES	(3) Naive Logit *ClickDummy*	(4) Logit with controls *ClickDummy*
ShopRating		-0.321^{***}
		(0.0137)
ShopGradesCount		$-5.96\text{e-}07^{***}$
		(5.57e-08)
PriceDifFromAvg		-3.319^{***}
		(0.295)
PriceDifFromMin		-1.850^{***}
		(0.560)
PriceDifFromDefault		-1.519^{***}
		(0.489)
PriceDifFromPremium		-2.546^{***}
		(0.143)
SessionLength		-0.000190^{***}
		(1.32e-05)
ModelRating		0.529^{**}
		(0.218)
ModelReviewsCount		0.000222
		(5.46e-05)
ModelFeature1 (AutonomousDuration)		-0.623^{***}
		(0.116)
ModelFeature2 (MemoryVolume)		0.236^{*}
		(0.131)
ModelFeature3 (Camera)		-0.438
		(0.306)
ModelFeature4 (Performance)		0.514^{**}
		(0.236)
ModelFeature5 (Screen)		-3.074^{***}
		(0.449)
Observations	1,244,191	1,244,191
Controls	no	yes
R2	0.0012	0.0511

*Note: Outcome variable is a dummy indicating whether the click on viewed offer happened. All variables and are rigorously described in Section 5. Clustered (by user sessions) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

9.5 Estimation Results: Model 5 for clicks.

This subsection presents the full results of Multinomial Logit estimation (Model 5) with click dummy being a dependent variable. The output is produced by *mlogit* command in Stata. Coefficients are estimated for each brand, with numbers 1 to 5 denoting Apple, Samsung, Xiaomi, Honor, and Huawei respectively. Outside option (0) is taken as the base outcome.

Table A.6: Full estimation results (for clicks), Model 5 (Multinomial Logit)

```
Multinomial logistic regression                 Number of obs     =  1,245,546
                                                Wald chi2(115)    =  108036.31
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -533700.15               Pseudo R2         =     0.2776

                             (Std. Err. adjusted for 247,727 clusters in visitid)
```

clickbranddummy	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
0	(base outcome)					
1						
offerprice	.0002136	6.07e-06	35.22	0.000	.0002017	.0002255
hasseenavgpriceproxyflag	.8456836	.4301646	1.97	0.049	.0025765	1.688791
bellshapedrop	.3717934	.238949	1.56	0.120	-.096538	.8401247
bellshapedropx	-.5296856	.4434861	-1.19	0.232	-1.398902	.3395312
ushaperise	-.0712598	.2302889	-0.31	0.757	-.5226178	.3800982
ushaperisex	-.2929804	.4358602	-0.67	0.501	-1.147251	.5612899
persistentflat	.3450847	.8580937	0.40	0.688	-1.336748	2.026917
persistentflatx	-14.20841	.8646296	-16.43	0.000	-15.90306	-12.51377
persistentfall	.2535299	.2384218	1.06	0.288	-.2137684	.7208281
persistentfallx	-.2821773	.4335683	-0.65	0.515	-1.131956	.567601
shoprating	-.1598157	.0102979	-15.52	0.000	-.1799992	-.1396321
shopgradescount	5.12e-08	4.38e-08	1.17	0.242	-3.46e-08	1.37e-07
pricediffromavg	-5.009649	1.668789	-3.00	0.003	-8.280415	-1.738882
pricediffrommin	-7.274296	2.088765	-3.48	0.000	-11.3682	-3.180393
pricediffromdefault	.1929945	.2831148	0.68	0.495	-.3619004	.7478893
pricediffrompremium	-.6754198	.4811554	-1.40	0.160	-1.618467	.2676274
sessionlength	.0001159	.0000229	5.06	0.000	.000071	.0001608
modelreviewscount	-.0026119	.0002505	-10.42	0.000	-.003103	-.0021208
modelfeature1autdur	16.91476	.6181045	27.37	0.000	15.7033	18.12622
modelfeature2memvol	13.84777	.3999951	34.62	0.000	13.0638	14.63175
modelfeature3camera	-9.333661	1.651209	-5.65	0.000	-12.56997	-6.09735
modelfeature4performance	-9.804918	1.259643	-7.78	0.000	-12.27377	-7.336062
modelfeature5screen	48.62466	1.764634	27.56	0.000	45.16605	52.08328
_cons	-286.5811	9.625201	-29.77	0.000	-305.4461	-267.716
2						
offerprice	-.0000758	.0000287	-2.64	0.008	-.0001321	-.0000195
hasseenavgpriceproxyflag	.8135676	.1135351	7.17	0.000	.5910429	1.036092
bellshapedrop	.0507018	.1455466	0.35	0.728	-.2345642	.3359679
bellshapedropx	-.3215116	.272982	-1.18	0.239	-.8565465	.2135234
ushaperise	-.3198471	.07626	-4.19	0.000	-.469314	-.1703802
ushaperisex	-.0574668	.1740274	-0.33	0.741	-.3985543	.2836207
persistentflat	.5142863	.2693538	1.91	0.056	-.0136375	1.04221
persistentflatx	-.7012651	.2495874	-2.81	0.005	-1.190447	-.2120828
persistentfall	-1.217221	.3679423	-3.31	0.001	-1.938375	-.4960673
persistentfallx	-.2323525	.3672527	-0.63	0.527	-.9521547	.4874496
shoprating	-.0966659	.0222441	-4.35	0.000	-.1402636	-.0530683
shopgradescount	-1.56e-07	6.80e-08	-2.30	0.022	-2.89e-07	-2.29e-08
pricediffromavg	-.5499216	.8132714	-0.68	0.499	-2.143904	1.044061
pricediffrommin	2.902322	.2947042	9.85	0.000	2.324712	3.479932
pricediffromdefault	.3995564	.2511565	1.59	0.112	-.0927014	.8918141
pricediffrompremium	-1.108012	.5609683	-1.98	0.048	-2.20749	-.0085342
sessionlength	-.0001973	.0000474	-4.16	0.000	-.0002902	-.0001043
modelreviewscount	-.0152676	.0012926	-11.81	0.000	-.017801	-.0127341
modelfeature1autdur	15.39099	.834404	18.45	0.000	13.75559	17.02639
modelfeature2memvol	51.04869	3.103743	16.45	0.000	44.96547	57.13192
modelfeature3camera	-26.11771	2.528242	-10.33	0.000	-31.07297	-21.16245
modelfeature4performance	-15.8311	2.090734	-7.57	0.000	-19.92887	-11.73334
modelfeature5screen	-2.108875	1.357335	-1.55	0.120	-4.769203	.5514524
_cons	-111.5074	11.05988	-10.08	0.000	-133.1844	-89.83045

3						
offerprice	-.0001604	8.07e-06	-19.87	0.000	-.0001762	-.0001446
hasseenavgpriceproxyflag	.5997856	.0798284	7.51	0.000	.4433248	.7562463
bellshapedrop	.2118292	.0568737	3.72	0.000	.1003587	.3232997
bellshapedropx	-.0382119	.1175869	-0.32	0.745	-.2686779	.1922541
ushaperise	.0328131	.045458	0.72	0.470	-.056283	.1219092
ushaperisex	.0282201	.0976557	0.29	0.773	-.1631815	.2196216
persistentflat	.5632267	.4256614	1.32	0.186	-.2710543	1.397508
persistentflatx	1.39654	.4472452	3.12	0.002	.5199555	2.273124
persistentfall	.0413575	.0521505	0.79	0.428	-.0608556	.1435706
persistentfallx	.2318852	.1004019	2.31	0.021	.0351011	.4286693
shoprating	-.0883422	.0157404	-5.61	0.000	-.1191928	-.0574916
shopgradescount	-3.08e-07	4.85e-08	-6.35	0.000	-4.03e-07	-2.13e-07
pricediffromavg	1.456081	.387575	3.76	0.000	.696448	2.215714
pricediffrommin	.9428853	.2217561	4.25	0.000	.5082515	1.377519
pricediffromdefault	.4116332	.1619005	2.54	0.011	.0943141	.7289523
pricediffrompremium	.5859167	.2754097	2.13	0.033	.0461235	1.12571
sessionlength	-.00001	.000019	-0.53	0.598	-.0000473	.0000273
modelreviewscount	-.0010535	.0001404	-7.51	0.000	-.0013286	-.0007784
modelfeature1autdur	5.899023	.4584159	12.87	0.000	5.000544	6.797501
modelfeature2memvol	8.059577	1.073107	7.51	0.000	5.956327	10.16283
modelfeature3camera	27.00412	1.947421	13.87	0.000	23.18724	30.82099
modelfeature4performance	-.8212179	.6243386	-1.32	0.188	-2.044899	.4024633
modelfeature5screen	16.80634	1.001053	16.79	0.000	14.84431	18.76837
_cons	-272.0605	10.56932	-25.74	0.000	-292.776	-251.345
4						
offerprice	2.95e-06	3.72e-06	0.79	0.429	-4.35e-06	.0000102
hasseenavgpriceproxyflag	.9631817	.1512123	6.37	0.000	.6668111	1.259552
bellshapedrop	.3744206	.0874908	4.28	0.000	.2029418	.5458995
bellshapedropx	-.5444382	.2089177	-2.61	0.009	-.9539095	-.134967
ushaperise	-.8694175	.1053242	-8.25	0.000	-1.075849	-.6629858
ushaperisex	-.8182773	.2583067	-3.17	0.002	-1.324549	-.3120055
persistentflat	-5.927125	.3774436	-15.70	0.000	-6.666901	-5.187349
persistentflatx	-.0240088	.6550743	-0.04	0.971	-1.307931	1.259913
persistentfall	.1974943	.0898347	2.20	0.028	.0214216	.373567
persistentfallx	-.5032063	.173365	-2.90	0.004	-.8429955	-.1634171
shoprating	-.2459079	.0229228	-10.73	0.000	-.2908358	-.2009801
shopgradescount	7.76e-07	6.65e-08	11.68	0.000	6.46e-07	9.06e-07
pricediffromavg	9.482874	.8576617	11.06	0.000	7.801888	11.16386
pricediffrommin	-10.27822	.7815743	-13.15	0.000	-11.81008	-8.746366
pricediffromdefault	.1347929	.1657739	0.81	0.416	-.1901181	.4597038
pricediffrompremium	2.133758	.4178205	5.11	0.000	1.314845	2.952671
sessionlength	-.0001248	.0000407	-3.07	0.002	-.0002046	-.0000451
modelreviewscount	-.0028303	.0002045	-13.84	0.000	-.003231	-.0024295
modelfeature1autdur	-4.055654	.3759535	-10.79	0.000	-4.792509	-3.318798
modelfeature2memvol	3.112688	.4935594	6.31	0.000	2.145329	4.080046
modelfeature3camera	-16.6412	1.175566	-14.16	0.000	-18.94527	-14.33713
modelfeature4performance	19.12571	.9826577	19.46	0.000	17.19973	21.05168
modelfeature5screen	-24.66595	.8051393	-30.64	0.000	-26.24399	-23.0879
_cons	106.2672	5.062383	20.99	0.000	96.34512	116.1893
5						
offerprice	.0000145	5.36e-06	2.70	0.007	3.96e-06	.000025
hasseenavgpriceproxyflag	.3839821	.0670373	5.73	0.000	.2525914	.5153729
bellshapedrop	.0055012	.0489107	0.11	0.910	-.0903621	.1013645
bellshapedropx	.1482905	.1207221	1.23	0.219	-.0883204	.3849015
ushaperise	-.2513211	.0509812	-4.93	0.000	-.3512424	-.1513997
ushaperisex	.1210435	.1387083	0.87	0.383	-.1508199	.3929068
persistentflat	.2726867	.513434	0.53	0.595	-.7336254	1.278999
persistentflatx	-16.57601	.4177055	-39.68	0.000	-17.39469	-15.75732
persistentfall	-.0694748	.1439828	-0.48	0.629	-.351676	.2127264
persistentfallx	.1465433	.273898	0.54	0.593	-.3902869	.6833736
shoprating	-.0854551	.0169087	-5.05	0.000	-.1185956	-.0523145
shopgradescount	-1.87e-07	4.08e-08	-4.59	0.000	-2.67e-07	-1.07e-07
pricediffromavg	2.396184	.6481039	3.70	0.000	1.125924	3.666445
pricediffrommin	-2.806813	.5431359	-5.17	0.000	-3.87134	-1.742286
pricediffromdefault	.6812214	.2621882	2.60	0.009	.1673419	1.195101
pricediffrompremium	.1227446	.260161	0.47	0.637	-.3871616	.6326507
sessionlength	-.0000289	.0000238	-1.21	0.226	-.0000756	.0000178
modelreviewscount	.0048692	.0004423	11.01	0.000	.0040023	.0057361
modelfeature1autdur	-1.461929	.2697147	-5.42	0.000	-1.99056	-.9332981
modelfeature2memvol	-7.271495	1.237395	-5.88	0.000	-9.696744	-4.846246
modelfeature3camera	-26.86148	.8691763	-30.90	0.000	-28.56504	-25.15793
modelfeature4performance	-7.491084	.4797643	-15.61	0.000	-8.431405	-6.550764
modelfeature5screen	39.91168	2.13373	18.71	0.000	35.72965	44.09372
_cons	3.962283	4.429155	0.89	0.371	-4.718701	12.64327

9.6 Estimation Results: Model 6 for clicks.

This subsection presents the full results of Nested Logit estimation (Model 6) with click dummy being a dependent variable. The output is produced by *mlogit* command in R with parameter *nests* included. Intercepts are brand specific, with model id shown as an identificator.

Apple model ids: 558168089, 175944418, 558163101

Samsung model ids: 650869003, 652152122, 521439130, 653533063, 661312003

Xiaomi model ids: 572745038, 662539009, 678180055, 637388421, 665736013, 663190021

Honor model ids: 417636374, 612791068, 453404067, 676859009, 662491009

Huawei model ids: 657023057, 657843127, 544612019, 418964205

Table A.7: Full estimation results (for clicks), Model 6 (Nested Logit)

```
Coefficients :
                          Estimate  Std. Error   z-value  Pr(>|z|)
(Intercept):175944418     1.0802e+01  9.5551e-02  113.0490 < 2.2e-16 ***
(Intercept):417636374    -2.9815e+00  1.0173e-01  -29.3085 < 2.2e-16 ***
(Intercept):418964205     9.8554e+00  8.9279e-02  110.3887 < 2.2e-16 ***
(Intercept):453404067     3.9324e-01  1.0321e-01    3.8100 0.0001389 ***
(Intercept):521439130     1.1923e+01  1.5266e-01   78.1029 < 2.2e-16 ***
(Intercept):544612019     1.6647e+00  6.3738e-02   26.1173 < 2.2e-16 ***
(Intercept):558163101     2.0162e+01  1.4208e-01  141.9103 < 2.2e-16 ***
(Intercept):558168089     1.5809e+01  1.2031e-01  131.3953 < 2.2e-16 ***
(Intercept):572745038     1.1625e+00  4.1916e-02   27.7337 < 2.2e-16 ***
(Intercept):612791068    -5.1249e-01  4.2335e-02  -12.1056 < 2.2e-16 ***
(Intercept):637388421     7.0586e-01  7.0773e-01    0.9974 0.3185893
(Intercept):650869003     1.6293e+00  4.3117e-02   37.7875 < 2.2e-16 ***
(Intercept):652152122     3.7141e+00  5.1032e-02   72.7802 < 2.2e-16 ***
(Intercept):653533063     1.3487e+01  1.0927e-01  123.4300 < 2.2e-16 ***
(Intercept):657023057    -1.2675e-01  5.5970e-02   -2.2646 0.0235395 *
(Intercept):657843127     1.3949e+01  1.1410e-01  122.2530 < 2.2e-16 ***
(Intercept):661312003    -1.0262e+00  5.0089e-02  -20.4878 < 2.2e-16 ***
(Intercept):662491009    -2.7017e+00  5.4816e-02  -49.2863 < 2.2e-16 ***
(Intercept):662539009     2.6281e+00  4.3223e-02   60.8035 < 2.2e-16 ***
(Intercept):663190021     2.8297e+00  6.5802e-02   43.0031 < 2.2e-16 ***
(Intercept):665736013     8.5634e-01  8.6682e-02    9.8790 < 2.2e-16 ***
(Intercept):676859009    -6.6391e-01  7.8660e-02   -8.4402 < 2.2e-16 ***
(Intercept):678180055     2.2198e+00  5.7768e-02   38.4258 < 2.2e-16 ***
OfferPrice               -3.2905e-04  2.0015e-06 -164.4007 < 2.2e-16 ***
HasSeenAvgPriceProxyFlag  3.1152e-01  4.1240e-02    7.5539 4.219e-14 ***
BellShapeDrop             1.3900e-01  2.6374e-02    5.2701 1.363e-07 ***
BellShapeDropX           -3.1972e-02  6.5037e-02   -0.4916 0.6230092
UShapeRise               -1.7156e-01  2.4435e-02   -7.0210 2.203e-12 ***
UShapeRiseX               1.8124e-01  5.7991e-02    3.1254 0.0017757 **
PersistentFall           -2.6075e-01  3.0262e-02   -8.6164 < 2.2e-16 ***
PersistentFallX           1.7666e-01  5.5977e-02    3.1560 0.0015997 **
SessionLength            -2.0859e-04  9.4557e-06  -22.0603 < 2.2e-16 ***
iv                        9.9376e-01  1.1761e-02   84.4944 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log-Likelihood: -121330
McFadden R^2:  0.060834
Likelihood ratio test : chisq = 16245 (p.value = < 2.22e-16)
```

9.7 Estimation Results: Models 1-2 for purchases.

Table A.8: Estimation results (for purchases), Models 1-2

	(1)	(2)
	Naive OLS	OLS with controls
VARIABLES	*Purchases*	*Purchases*
BellShapeDrop	0.386***	0.400***
	(0.042)	(0.043)
UShapeRise	0.377***	0.392***
	(0.047)	(0.050)
PersistentFlat	−0.293***	−0.302***
	(0.041)	(0.042)
PersistentFall	0.532***	0.520***
	(0.078)	(0.081)
TotalPurchases	8.93e-06***	9.69e-06***
	(3.43e-07)	(2.88e-07)
Price_1000-3000_Dummy	0.407***	0.302***
	(0.030)	(0.028)
Price_3000-10000_Dummy	0.797***	0.665***
	(0.048)	(0.042)
Price_10000-25000_Dummy	0.780***	0.548***
	(0.055)	(0.044)
Price_25000+_Dummy	0.268***	−0.137***
	(0.027)	(0.014)
OfferDiffFromAvgPrice (ODAvg)		−0.0037**
		(0.001)
ShopRating		0.433***
		(0.010)
Constant	−0.364***	−1.264***
	(0.071)	(0.076)
Observations	2,255,922	2,255,922
Product category dummies	no	yes
R2	0.003	0.014

*Note: Outcome variable is a number of purchases made in particular group. Among dummies for price buckets, the omitted one is [Price 0-1000 Dummy]. All employed variables are rigorously described in Section 5. Clustered (by user sessions) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

9.8 Estimation Results: Models 3-5 for purchases.

Table A.9: Estimation results (for purchases), Models 3-5

VARIABLES	(3) Naive Logit *PurchaseDummy*	(4) Logit with controls *PurchaseDummy*	(5) Multinomial Logit[1] *PurchaseBrandDummy*
OfferPrice	−1.84e-05***	−3.40e-05***	−1.65e-05***
	(1.99e-06)	(8.94e-06)	(2.51e-06)
DSH	0.301	0.356*	0.866***
	(0.192)	(0.194)	(0.227)
BellShapeDrop	0.400***	0.240*	0.142
	(0.115)	(0.123)	(0.168)
BellShapeDrop ×	0.379	0.356	−0.094
	(0.269)	(0.269)	(0.302)
UShapeRise	0.351***	0.136	−0.762***
	(0.104)	(0.115)	(0.184)
UShapeRise ×	0.139	0.142	−2.126***
	(0.246)	(0.246)	(0.562)
PersistentFlat	−1.589***	−2.411***	−1.764***
	(0.583)	(0.676)	(0.396)
PersistentFlat ×	0.550	0.462	
	(1.170)	(1.164)	
PersistentFall	0.132	−0.284*	−1.072***
	(0.113)	(0.145)	(0.169)
PersistentFall ×	0.181	−0.187	−0.250
	(0.259)	(0.261)	(0.286)
Constant	−6.915***	−2.291	−29.840***
	(0.0925)	(13.20)	(3.314)
Observations	1,244,191	1,244,191	1,244,217[2]
Controls	no	yes	yes
R2	0.0120	0.0609	0.1166

*Note: Outcome variable is a dummy indicating whether the purchase of viewed offer happened. For Multinomial Logit in case of purchase dummy takes the value of purchased brand index. All variables and are rigorously described in Section 5. Mark × denotes the interaction with DSH dummy. Clustered (by user sessions) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

[1] *Some controls were dropped due to multicollinearity issues. Results are presented only for one selected group: Honor-branded phones. For full results see Table A.11 in Appendix 9.9.*

[2] *The outcome variable were tracked and extrapolated through the same session. In case of rare yet existing successfull outcomes for multiple brands during one session this required artificial session duplication.*

Table A.10: Estimation results (for purchases) for control variables, Models 3-4

VARIABLES	(3) Naive Logit *PurchaseDummy*	(4) Logit with controls *PurchaseDummy*
ShopRating		-0.626^{***}
		(0.0595)
ShopGradesCount		$-9.71\text{e-}07^{***}$
		(2.96e-07)
PriceDifFromAvg		-0.471
		(1.057)
PriceDifFromMin		-2.846
		(1.998)
PriceDifFromDefault		-5.440^{***}
		(1.774)
PriceDifFromPremium		-2.412^{***}
		(0.617)
SessionLength		$-2.33\text{e-}05$
		(4.17e-05)
ModelRating		-0.478
		(0.831)
ModelReviewsCount		0.000181
		(0.000243)
ModelFeature1 (AutonomousDuration)		-1.640^{***}
		(0.570)
ModelFeature2 (MemoryVolume)		0.254
		(0.705)
ModelFeature3 (Camera)		3.111^{*}
		(1.684)
ModelFeature4 (Performance)		2.765^{*}
		(1.482)
ModelFeature5 (Screen)		-1.952
		(2.365)
Observations	1,244,191	1,244,191
Controls	no	yes
R2	0.0120	0.0609

*Note: Outcome variable is a dummy indicating whether the purchase of viewed offer happened. All variables and are rigorously described in Section 5. Clustered (by user sessions) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

9.9 Estimation Results: Model 5 for purchases.

This subsection presents the full results of Multinomial Logit estimation (Model 5) with the purchase dummy being a dependent variable. The output is produced by *mlogit* command in Stata.[68] Coefficients are estimated for each brand, with numbers 1 to 5 denoting Apple, Samsung, Xiaomi, Honor, and Huawei respectively. Outside option (0) is taken as the base outcome.

Table A.11: Full estimation results (for purchases), Model 5 (Multinomial Logit)

Multinomial logistic regression

Number of obs	=	1,244,217
LR chi2(80)	=	13832.63
Prob > chi2	=	0.0000
Pseudo R2	=	0.1166

Log likelihood = -52382.253

purchasebranddummy	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
0	(base outcome)					
1						
offerprice	.0000418	1.67e-06	25.01	0.000	.0000385	.000045
hasseenavgpriceproxyflag	-.0850592	1628.642	-0.00	1.000	-3192.165	3191.995
bellshapedrop	17.00149	563.954	0.03	0.976	-1088.328	1122.331
bellshapedropx	-.6378376	1628.642	-0.00	1.000	-3192.718	3191.442
ushaperise	17.06836	563.954	0.03	0.976	-1088.261	1122.398
ushaperisex	.6911576	1628.642	0.00	1.000	-3191.389	3192.771
persistentflat	-.5097784	2148.629	-0.00	1.000	-4211.746	4210.726
persistentflatx	-.179113	6103.722	-0.00	1.000	-11963.25	11962.9
persistentfall	17.18647	563.954	0.03	0.976	-1088.143	1122.516
persistentfallx	.3268631	1628.642	0.00	1.000	-3191.753	3192.407
shoprating	-.2787015	.0495733	-5.62	0.000	-.3758633	-.1815397
shopgradescount	6.50e-07	1.75e-07	3.71	0.000	3.07e-07	9.93e-07
pricediffromavg	6.803183	.959844	7.09	0.000	4.921923	8.684442
pricediffrommin	-9.332375	.8302679	-11.24	0.000	-10.95967	-7.70508
sessionlength	.0001292	.0000299	4.31	0.000	.0000705	.0001879
modelfeature1autdur	2.963195	.5249037	5.65	0.000	1.934402	3.991987
_cons	-34.8473	563.9597	-0.06	0.951	-1140.188	1070.493
2						
offerprice	-.0002188	.0000119	-18.45	0.000	-.000242	-.0001956
hasseenavgpriceproxyflag	.4141691	.2589449	1.60	0.110	-.0933535	.9216918
bellshapedrop	.37189	.184006	2.02	0.043	.011245	.7325351
bellshapedropx	.5088152	.3833363	1.33	0.184	-.2425101	1.260141
ushaperise	1.186923	.1265848	9.38	0.000	.9388218	1.435025
ushaperisex	-.2260295	.3133865	-0.72	0.471	-.8402558	.3881969
persistentflat	1.648857	.192113	8.58	0.000	1.272322	2.025391
persistentflatx	.0198962	.4312188	0.05	0.963	-.8252771	.8650696
persistentfall	-.6894811	.2876076	-2.40	0.017	-1.253182	-.1257807
persistentfallx	-14.88395	889.0177	-0.02	0.987	-1757.327	1727.559
shoprating	-.6575522	.0820091	-8.02	0.000	-.8182871	-.4968172
shopgradescount	-5.44e-07	3.13e-07	-1.74	0.083	-1.16e-06	7.01e-08
pricediffromavg	.977744	.5748864	1.70	0.089	-.1490127	2.104501
pricediffrommin	1.751652	.1139811	15.37	0.000	1.528253	1.975051
sessionlength	-.0001414	.0000569	-2.49	0.013	-.0002529	-.0000299
modelfeature1autdur	.2315956	.5152902	0.45	0.653	-.7783545	1.241546
_cons	-4.689091	2.498983	-1.88	0.061	-9.587007	.2088257
3						
offerprice	-.0000807	2.40e-06	-33.65	0.000	-.0000854	-.000076
hasseenavgpriceproxyflag	.6473898	.126035	5.14	0.000	.4003657	.8944139
bellshapedrop	1.182847	.0752196	15.73	0.000	1.03542	1.330275
bellshapedropx	.3114643	.1486977	2.09	0.036	.0200222	.6029064
ushaperise	1.1662	.0705666	16.53	0.000	1.027892	1.304508
ushaperisex	.1913798	.1414868	1.35	0.176	-.0859292	.4686889
persistentflat	-17.88549	1518.47	-0.01	0.991	-2994.031	2958.26
persistentflatx	-.7118406	4290.871	-0.00	1.000	-8410.664	8409.24
persistentfall	.6896056	.0793679	8.69	0.000	.5340474	.8451638
persistentfallx	.4922699	.1434929	3.43	0.001	.211029	.7735109
shoprating	.3215358	.0396512	8.11	0.000	.2438208	.3992507
shopgradescount	-4.36e-07	1.27e-07	-3.42	0.001	-6.86e-07	-1.86e-07
pricediffromavg	7.33865	.4968579	14.77	0.000	6.364826	8.312473
pricediffrommin	-4.571174	.3900681	-11.72	0.000	-5.335694	-3.806655
sessionlength	.0000936	.0000167	5.62	0.000	.0000609	.0001262
modelfeature1autdur	4.5555	.2338257	19.48	0.000	4.09721	5.01379
_cons	-29.6066	1.186135	-24.96	0.000	-31.93139	-27.28182

[68]The standard errors are questionable since we were unable to identify a model with SEs clustered on the level of visit.

4						
offerprice	-.0000165	2.51e-06	-6.58	0.000	-.0000214	-.0000116
hasseenavgpriceproxyflag	.8658543	.2271646	3.81	0.000	.4206199	1.311089
bellshapedrop	.1415058	.1684822	0.84	0.401	-.1887132	.4717248
bellshapedropx	-.093972	.3016222	-0.31	0.755	-.6851406	.4971966
ushaperise	-.7616633	.184278	-4.13	0.000	-1.122842	-.400485
ushaperisex	-2.126122	.5620314	-3.78	0.000	-3.227683	-1.024561
persistentflat	-1.764002	.3962377	-4.45	0.000	-2.540613	-.9873898
persistentflatx	-19.54727	8737.275	-0.00	0.998	-17144.29	17105.2
persistentfall	-1.072286	.1694574	-6.33	0.000	-1.404416	-.7401554
persistentfallx	-.2495978	.2861623	-0.87	0.383	-.8104656	.3112699
shoprating	-.5718399	.080207	-7.13	0.000	-.7290426	-.4146371
shopgradescount	7.86e-07	3.10e-07	2.53	0.011	1.78e-07	1.39e-06
pricediffromavg	25.29994	1.229439	20.58	0.000	22.89028	27.70959
pricediffrommin	-22.13278	1.05559	-20.97	0.000	-24.2017	-20.06386
sessionlength	.0002475	.0000358	6.91	0.000	.0001773	.0003177
modelfeature1autdur	5.368406	.6766562	7.93	0.000	4.042184	6.694628
_cons	-29.84012	3.313796	-9.00	0.000	-36.33504	-23.3452
5						
offerprice	-.0000476	1.87e-06	-25.44	0.000	-.0000513	-.000044
hasseenavgpriceproxyflag	.6162902	.0766359	8.04	0.000	.4660866	.7664937
bellshapedrop	-.0508545	.0638599	-0.80	0.426	-.1760175	.0743085
bellshapedropx	.6415087	.1214666	5.28	0.000	.4034386	.8795788
ushaperise	-.5165256	.06613	-7.81	0.000	-.6461381	-.3869132
ushaperisex	.385093	.132447	2.91	0.004	.1255017	.6446843
persistentflat	-19.75458	1833.764	-0.01	0.991	-3613.866	3574.357
persistentflatx	-.5734498	5164.914	-0.00	1.000	-10123.62	10122.47
persistentfall	-3.913084	.46167	-8.48	0.000	-4.817941	-3.008228
persistentfallx	-14.06491	875.0942	-0.02	0.987	-1729.218	1701.088
shoprating	.0665699	.048808	1.36	0.173	-.029092	.1622318
shopgradescount	-1.29e-07	1.50e-07	-0.86	0.388	-4.23e-07	1.64e-07
pricediffromavg	.9518784	.1474251	6.46	0.000	.6629306	1.240826
pricediffrommin	.6994727	.0826904	8.46	0.000	.5374024	.861543
sessionlength	.0001253	.0000206	6.10	0.000	.000085	.0001656
modelfeature1autdur	-6.000377	.1557429	-38.53	0.000	-6.305628	-5.695127
_cons	20.74413	.7732686	26.83	0.000	19.22856	22.25971