No More Free Choice: An Empirical Analysis of Default Options in the Amazon's Online Marketplace

Guido Deiana

Thesis Supervisor: Professor Francesco Decarolis

Università Bocconi

Abstract

The purpose of this work is to estimate the impact on consumers' demand of products suggestions in an online marketplace. To do so, a discrete choice model of consumer demand is developed in two variations; the first one following the Industrial Organization literature and the second one accounting for bounded rationality and behavioral biases in customers. The model is then estimated with modern Empirical Industrial Organization methods, among which the Nested Logit and the Random Coefficient Logit (BLP) are used. The setting proposed is the Amazon marketplace and specifically this work studies the role of the Amazon's Choice badge in influencing consumers' demand. It is shown that the badge has a strong and statistically significant effect across the different structural models and the various empirical techniques implemented. These results show the potential for distortion of consumers' demand coming from a vertically integrated platform which competes with third party sellers in its own environment.

Contents

1 Introduction

In the last twenty years, the rise in popularity of online shopping has shown that online marketplaces can be a viable alternative to physical stores. The internet provides users with a quick access to a large variety of products, fast shipping and often cheaper prices than physical shops. In the last few years especially, there have been major improvements in the speed of shipping and in the variety of products available. However, the rise in concentration of online marketplaces (among them Ebay in the early 2000s and now Amazon) is a clear sign that large systems of quality assurance, seller reputation and shipping management have major fixed costs that need a large customer base to be implemented. Complex methods of platform personalization, obfuscation of competitors' prices and non-apparent switching costs have been implemented in the last years by many of these platforms in order to secure customers and lock them in a single integrated environment.

When online shopping first appeared in the early days of the internet, it was welcomed as a highly competitive marketplace with virtually null search costs (any product is just a few clicks away). It was difficult to foresee then the competitive spectrum that can be seen today, with a highly concentrated market in the hands of very few, heavily integrated, successful platforms that act as intermediaries between buyers and sellers. The recent competitive concerns on the power of such giants, raised both by the European Commission in Europe and by the Federal Trade Commission in the United States, show that there are ways in which these platforms could gain advantages against other competitor platforms, or against individual online sellers.[1]

The purpose of this work is to investigate one specific way in which online platforms could be influencing consumers' choice. More specifically, this work builds on the assumption that buyers in online marketplaces are often in a situation called "Choice Overload". As explained by Spiegler (2011): "The larger the set of market alternatives, the greater consumers' ten-

[1] On July 17, 2019 the European Commission opened a formal investigation to assess whether the use of sellers' data from Amazon is a breach of EU competition rules. See: European Commission (2019)

dency to evade the choice problem and adhere to the default". In this setting, the abundance of similar products available to the buyer can be exploited by the platform itself to suggest or indicate a default product which customers tend to buy to avoid having to compare all the other possible options.

Default choices have been long addressed by the Economics literature since the rise in popularity of Behavioral Economics. However, the setting analyzed is more subtle than an explicit default choice and requires a mix of traditional Industrial Organization discrete choice models of demand and behavioral adaptations of such models. This work focuses on the Amazon Marketplace and on the recently introduced Amazon's Choice product badge.[2] The badge appears in almost every product research made by customers and identifies one specific product as the one suggested by Amazon for the specific combination of words used by the customer in the research (it is important to note that very similar researches could be associated to different Amazon's Choice products). The Amazon's Choice is therefore not a default option strictly speaking, at least not when shopping thorough the website, but rather a strong suggestion that the website proposes to customers together with other products.

One of the main issues with the "Amazon's Choice" badge is that it is that it is unclear, to buyers and sellers alike, how the badge is assigned to products. It has been speculated that it is likely to be assigned to products highly available, quick to ship to customers and with a high ratio of positive reviews. Nonetheless, in these years Amazon has never given a definitive answer on how the "Amazon's Choice" badge was assigned.[3] The validity of the badge as a form of quality assurance has also been questioned by the press (Shifflett et al., 2019), with Amazon never replying. Complicating things further, Amazon participates in the upstream market as a producer and a seller, competing against other third party independent sellers in its own platform. This particular setting therefore raises concerns regarding the

[2]The Amazon's Choice badge was first introduced to simplify the customers' shopping experience through voice assistants (i.e: Alexa) which would directly buy the Amazon's Choice product by default. It was then immediately integrated in the Amazon website as a non-intrusive badge on top of the product's image

[3]The only information provided by Amazon on the topic is a message box that reports: "Amazon's Choice recommends highly rated and well-priced products."

influence on consumer's choices that Amazon can exert in its own platform.

This work aims at quantifying the impact of the Amazon's Choice badge on consumers' demand. To do so, a discrete choice theoretical model is developed. In this model, the Amazon's Choice badge is considered a simple characteristic of the product and all the consumers are perfectly rational. Then, the model is further developed in a bounded rationality variation which considers two different types of customers; the first type (the rational consumer) considers all products and is not influenced by the Amazon's Choice badge, the second (the behavioral consumer) only considers the Amazon's Choice product and buys it as long as the expected utility of the product is larger than its reservation utility level. This last approach aims at capturing the heterogeneity of consumers in the choice mechanism, assuming that certain customers do not have enough time to compare all the options or that their attention is immediately captured by the Amazon's Choice badge, leading them to ignore all other options available on the platform.

For the empirical application, this work uses data on more than 1000 products divided in the Amazon Europe websites (Italy, France, Spain, Germany, UK) and in five subcategories of electronic cables (dvi, ethernet, hdmi, lightning, usb). Electronic cables are chosen for the following reasons: they have distinct and easily recognizable characteristics, both from the consumer's and from the researcher's standpoint; cables with the same characteristics are very close to being perfect substitutes; the mapping of the products in the Amazon subcategories is consistent with the cables end use, leading to a reliable market size estimation of cables sold on the Amazon marketplace. A web scraper is used to create the sample and gather static information about the products. An external source[4] is then used to gather historical sales and prices data for each of the selected products. From these sources, current data on products characteristics and Amazon's Choice badge is derived, as well as current and historical data on prices, sales, number of reviews and ratings of products.

The theoretical models of consumer choice described above are then estimated using

[4] `https://keepa.com/#!api`

econometric techniques from the Empirical Industrial Organization literature. Four different methodologies are used throughout the work. For the Full Rationality approach: the Multinomial Logit, the two-levels Nested Logit and the Random Coefficient Logit models (or BLP from S. Berry, Levinsohn, & Pakes (1995)) are used. For the Bounded Rationality approach, this work relies instead on a three-levels variation of the Nested Logit model. All models are estimated with and without brand fixed effects. Current endogenous variables (prices, number of reviews, ratings and Amazon's Choice badge) are instrumented with Arellano-Bond type instruments.

The Amazon's Choice badge shows a strong and statistically significant effect on consumer demand across all the models' variations. The order of magnitude of the Amazon's Choice badge is also strikingly large, showing a high potential for demand distortion by the platform. In fact, numerous channels through which the consumers decide to follow the Amazon's Choice badge are possible, some are analyzed here, but investigating them in depth would require more detailed (and possibly individual) data on consumers' choice. Such coefficients could be driven indeed by: time pressure or high value of time from the consumers' side; Amazon's Choice acting as an attention grabber, drawing consumers' attention towards the suggested product and away from the alternatives; Amazon's Choice badge being perceived by customers as an element of quality assurance directly issued by the platform.

The two theoretical variations of the discrete choice model presented above are developed in parallel due to the lack of evidence towards one channel or the other. The purpose of this work is in fact to quantify the impact of the Amazon's Choice badge, not to speculate on the channels through which it could be acting. With this premise, the two are presented as compelling evidence for the fact that the Amazon's Choice badge has a strong potential to distort consumers' choice. The evidence provided holds whether the value of the Amazon's Choice badge is overestimated by consumers (as in the fully rational consumer model), or whether the badge captures consumers' attention before they are able to compare all products (as in the behavioral variation).

These results have important policy implications in the fields of Competition Economics and Antitrust. The badge could in fact bring an unfair advantage to Amazon when it competes with other sellers upstream. So far its badge assignment has been an obscure mechanism which could be advantageous for certain sellers and discriminatory for others. This work also shows the potential for more general subtle demand distortions in online platforms, which could also happen in other (more or less effective) forms. Ultimately, the evidence presented supports the argument that in marketplaces with numerous similar products, when too many choices are available, customers can fail to optimize and often prefer following a generic default choice.

The reminder of this work is structured as follows. Section 2 covers and reviews the relevant literature for this work. Section 3 presents the main data sources and sampling techniques used. Section 4 presents the theoretical models of discrete choice estimated in the following sections. Section 5 and 6 show and explain the empirical models used and the identification strategies implemented. Section 7 covers the results of the empirical models presented in the previous sections. Section 8 draws the conclusions of the work. The appendix at the end covers the steps for the derivation of the three-levels Nested Logit model.

2 Literature Review

This work builds on the literature of traditional models of consumer behavior in Industrial Organization, as well as on more modern approaches to consumers' choices and consumers' attention. Together with this theoretical framework, an empirical approach is proposed, taken mostly from the modern applications of discrete choice methods in Empirical Industrial Organization. Ultimately, this work draws from the very recent publications on online markets and two sided platforms.

2.1 Bounded Rationality in Industrial Organization

In reviewing the relevant literature of bounded rationality in Industrial Organization, it is necessary to start with the work of both Ellison (2006) and Spiegler (2011) which provide an overview of modern applications of bounded rationality and Behavioral Economics models to the practice of Industrial Organization. The first paper by Ellison is more historically framed and lists the developments in behavioral Industrial Organization throughout the whole second half of the 20th century and the early years of the 21st. While the second work by Spiegler is directed more towards the use of traditional Industrial Organization models interacted with biases and insights from the recent studies in Behavioral Economics.

Ellison considers three important approaches in Behavioral Industrial Organization. The "Rule of Thumb" approach, in which consumers simply choose the first available option if it is satisfying, without comparing it to the other possibilities. The "Explicit Bounds" approach, in which consumers and firms are modelled to have explicit bounds in rationality and do not act as perfectly rational utility (or profit) maximizers. Lastly the "Psychology and Economics" approach, which draws from the literature following Tversky (1974) and uses the insights of psychology and behavioral economics to model consumer choice in an unconventional manner. All of these approaches are relevant and some of the assumptions deriving from these have been used by Industrial Organization economists well before the rise in popularity of behavioral economics. The models of Smallwood & Conlisk (1979) and Varian (1980) were in fact already considering imperfectly informed consumers, that choose which product to buy without considering all the possible options. Even before that, Radner (1975) was already considering agents with limited capacity to consider all the possible options available and settle for a "satisfycing" choice rather than a utility maximizing one.

Today, as shown in the large work by Spiegler (2011), the use of behavioral models has become a fundamental part of the Industrial Organization set of tools to model consumers behavior.

2.2 Consumer Attention

With the rise of internet giants, the role of consumer attention has become central in the Industrial Organization debate. Modern models such as Prat & Valletti (2019) consider consumer attention as a scarce resource used as a product from online platforms (referred to as “attention brokers”) to sell to advertising companies. Modern developments in Industrial Organization theory and Behavioral Economics have also shown how firms exploit behavioral biases to price their products. This work follows the literature of theoretical industrial organization models that assess the effect of such behavioral biases on consumer demands and how these can affect products competition in standard marketplaces. Regarding this, the work of Bordalo, Gennaioli, & Shleifer (2013, 2015) is remarkably useful in understanding the setting of this work. The authors use consumer attention as a limited resource which, according to the market setting, can be drawn to one salient characteristic of the product (in their proposed case this is either price or quality). Each consumer then puts disproportionately high emphasis on such characteristic, weighting the other much less in terms of its own utility and ultimately chooses almost only according to the salient characteristic itself.

Price dispersion has already been mentioned as a recurring topic in Industrial Organization (Varian, 1980). Where price dispersion was considered in the past as a tool used by sellers to differentiate among the various types of consumers (Stahl, 1989), more recent developments in Industrial Organization have further defined the sources of price dispersion. In Grubb (2015), it has been related to consumer confusion and failure to correctly assess the quality of products. A similar result is also shown by Bronnenberg et al. (2015) in the pharmaceuticals market, where uninformed consumers often fail to recognize that products are perfect substitutes and are willing to spend an extra amount of money just to buy the more popular brand. The brand markup is therefore wrongly perceived as an assurance of high quality.

This work uses the setting of bounded rationality and the insights from behavioral economics as the main theoretical explanation for the empirical results provided in the following sections on the consumer response to default choices. However, it is also necessary to mention the recent developments in two sided markets and online platforms specifically, to have a better understanding of the competitive framework of Amazon's Marketplace.

2.3 Two Sided Platforms and Internet Markets

Amazon works as a two sided platform by bringing together buyers and sellers. It also competes with sellers by selling itself certain products ("Sold by Amazon"), shipping them and/or producing them itself under its signature brands ("Amazon Basics", "Alexa", "Kindle"...).

This work does not analyze competition among multiple internet markets, but instead it focuses on competition and demand estimation within the single Amazon Marketplace. It is nevertheless important to mention some of the most important portions of the recent literature on two sided markets. Caillaud & Jullien (2003) is one of the first theoretical models that looked into competition among different platforms, showing the importance of frictions and barriers generated by network externalises. For platforms, it is difficult to attract sellers as they only go where customers are; while it is difficult to attract customers where there are not enough sellers. This "Chicken and Egg" problem is traditionally solved by a strategy of divide and conquer: offer a discount on one side (divide) while exploiting the other side (conquer). This model was then taken and expanded with variations in numerous other publications both for theoretical models (Rochet & Tirole, 2006; Armstrong, 2006) and empirical applications (Zervas, Proserpio, & Byers, 2017).

It is out of the scope of this work to try and understand the dynamics of online markets competition, however it is curios to see how the various attempts of some of the most important internet marketplaces to lock-in the customers in their own environments. Tools proposed by Amazon, such as "Amazon Prime" and "Amazon Pantry" are ways to lock in online customers, which were traditionally considered multi-homers (using more than one platform). This dynamic was already well understood by some economists in the early 2000s (Ellison & Ellison, 2005) where business-to-customer platforms were seeing a lot of concentration (that was the case of Ebay for example); while business-to-business and job market platforms still showed a high degree of dispersion. The more recent work of Eliaz & Spiegler (2011) describes the use of attention grabbers as bait products used by online platforms to lock in customers and sell other products from the same website.

This work looks at competition within the Amazon platform, more than at competition among different platforms, and specifically at how reputation of products and platform suggestions affect consumers' demand. Online marketplaces drastically decrease the asymmetric information situation between buyer and seller by providing feedback on the different sellers, this leads to positive consumer surplus effects and decrease in low quality sellers (Saeedi, 2014). With a system of sellers' reputation in place, the behavior of a potentially dishonest seller is put in check by the reviews that customers provide on the product they buy (Tadelis, 2016). Today Amazon provides two systems in place: reviews on products and sellers and, more importantly, a single product for each search highlighted as "Amazon's Choice", which works as a feedback from Amazon itself on the high quality of the product. Similar to this tool, Ebay applied a badge to verify sellers selling on its platform. This worked as a formal guarantee of quality issued from Ebay itself. Welfare impacts of the Ebay badges have been found to be positive and significant by Hui et al. (2016).

2.4 Discrete Choice Models in Empirical Industrial Organization

At last this work builds on the recent developments in Empirical Industrial Organization and specifically in discrete choice methods for demand estimation. In the next sections, two variations of the popular Logit model for demand estimation are used: respectively Nested Logit and Random Coefficient Logit models (often referred as BLP).

In the traditional Multinomial Logit model, under the assumption that each customer buys at most one unit of a single product, market shares are used as estimates of average probabilities that consumers will choose a certain product. With few other assumptions and little computational complexity, this allows to model the agents utility function in terms products characteristics and price. More complex and precise versions of the Multinomial Logit model have been developed in order to have less restrictive assumptions and higher estimation precision. The next sections build on the work of S. Berry, Levinsohn, & Pakes (1995) in demand estimation through Nested Logit and Random Coefficient methods. The application of the Random Coefficient method follows specifically two more recent applications of the method, respectively Nevo (2001) and Decarolis (2015). On the Nested Logit side, the work of Goldberg & Verboven (2001) is remarkably useful in explaining the scope and application of Nested logit models with more than two levels.

Recent works in Empirical Industrial Organizations which exploit behavioral biases from consumers and follow a similar approach to the one presented in this work are also worth mentioning. In Hortaçsu, Madanizadeh, & Puller (2017) an empirical model for consumer inertia in the Texas electricity market is developed, in which consumers don't optimize their utility every period. The source of this inertia are identified in consumer inattention and incumbent advantage. Another important discrete choice model with behavioral biases is applied in Brown & Jeon (2020), in which the authors develop a closed form solution for a model of consumer choice in the Insurance market. Consumer inattention is taken into account as some insurance characteristics are considered immediately available (such as the insurance premium), while other pieces of information require some effort to be extracted by the potential customer.

The work presented here follows a more traditional application of Random Coefficient Logit and Nested Logit models than the two papers cited previously. The theoretical modelling and the setting for consumer choice are, however, rather symilar to the ones presented in these recent papers.

3 Data

The dataset used in this work is composed of a subsample of Amazon's products. The main goal in samlpling the correct subset of products was finding categories of products which were largely substitutable and with few easily identifiable characteristics. To do so, the market for electronic cables is selected and 5 categories of cables (usb, hdmi, vga, lightning, dvi) are analyzed. Amazon products are all assigned to a macro-category and numerous sub-categories.

For this work, the 5 most populated subcategories of the "cables" category are used, all of which belong to the Electronics or Informatics macro-categories. The macro-categories are useful in the dataset construction as each product is assigned a "Best Seller's Rank" (BSR) which ranks all the products in a category and in a geographical market according to how much quantity they are selling in the current month. The BSR is of crucial importance in trying to derive a more reliable estimation of sales.

From each category the data on the 100 most sold products are gathered for the same 5 categories in the 5 Amazon Europe markets (Italy, France, Germany, Spain, United Kingdom). There are some products which are repeated among the different markets, but sometimes they display different characteristics (like the "Amazon's Choice" badge which is market specific) and prices.

3.1 Amazon Europe Scraped Data

A certain amount of data is collected through the use of a web scraper as it is not present in any available dataset. The scraper is used to collect the url of all the 100 most sold products in each cable subsection in the five European Amazon markets. Also the main element of this analysis, the Amazon's Choice badge, is collected through the use of a scraper which enters each product's page and collects whether the product displays an "Amazon's Choice" badge or not.

This is possible thanks to the fact that Amazon displays the badge not only in the research page but also in each product's individual page. Therefore by going into each product's page it is possible to know whether the product is assigned any type of badge ("Amazon's Chocie", "Most Sold",...).

3.2 Keepa API

For the time-series data collection, this work relies on the keepa api database[5]. Keepa is a website which tracks prices, BSR, reviews and average score of all products available on Amazon Marketplace as well as Ebay prices for comparison (when an ebay alternative seller is available). From Keepa, the historical data of all the analyzed products is downloaded and then the BSRs and the prices are averaged on a monthly basis to mitigate the impact of isolated data collection errors, which sometimes were found in less popular products.

3.2.1 Best Seller's Rank and Junglescout.com

In order to estimate the correct market shares, the BSRs of products are not enough as there isn't a linear relationship between BSR and monthly sales. However, since BSR is a measure of products' popularity, there are estimates of monthly sales available which are constructed from BSR, category and market. The ones used here come from the website Junglescout.[6] Other estimators tend to give similar estimates to the one of Junglescout, but this work refers only to Junglescout as it is the quickest and most reliable to work with. A web scraper id used here as well to derive an estimated function relating BSR with monthly sales.

Amazon does not directly reveal how the BSR is assigned, as for the Amazon-choice badge, however it is broadly defined as an index of the most sold products within a category. The BSR is calculated on the basis of recent and historical sales, with more weight on the recent. There is anecdotal evidence of this in the BSR time series of least popular products. These products often show a dip in BSR when there is a single sale and then, over time,

[5] `https://keepa.com/#!api`

[6] `https://www.junglescout.com/estimator/`

they return to their baseline (no sales) BSR level. Having found this, the approach used is to average the BSR on a monthly basis, and then calculate sales according to the average BSR, this should give a rough estimate of the sales of a product, without being influenced by sudden spikes in the sales. This approach has the potential issue of not estimating the correct market shares. This could lead to major problems for the Logit approach as the Logit model does not allow errors in the market share definition and can be highly biased when products display market shares close to zero. To mitigate this issue, all products with a small market share (threshold set at 0.05%) are dropped from the sample. These are the products more likely to create biases due to low market shares as well as being the ones more likely to have a relatively higher estimation error in their market share. With this approach, the aim is to eliminate the small market share bias, as well as mitigate the errors in market shares calculation.[7]

3.2.2 Prices

The time series data on prices is obtained from the keepa api as well. As keepa.com mainly works as a price tracker, it collects data not only on the Amazon official price, but also on alternative new product prices from third party sellers selling on Amazon, as well as Ebay prices (if an Ebay listing counterpart of the same product exists) and used product prices (if the same product is also sold as used through the Amazon platform).

This work only focuses on directly comparable products, therefore Ebay and used prices are dropped. These could have been used in the instrumental approach as instruments form Amazon prices, but unfortunately a used or Ebay counterpart is rarely available. The choice for the correct price indicator is therefore restricted to Amazon and new prices. To construct the best price indicator is better to first reconstruct the user experience when buying a product from Amazon. The platform in fact, if possible, always displays the Amazon price, relegating third party prices in a second ("Other Offers") page, even if this price could be

[7] An alternative solution for the issue of market shares calculated with an error term is proposed by Gandhi et al. (2013), however it has only been theorized and applied to simulated data.

lower than the Amazon price. If an Amazon price is not available, the platform displays the lowest and fastest to ship offer for the new product in the main page, relegating used products and other prices in the "Other Offers" page. For simplicity, in this work individuals are assumed to only select the first page product and never comparing different offers for the same product. A comparison of all possible prices for each product would require a much larger effort and a consideration of the supply side competition as well. To derive the prices displayed in the main product page at each point in time, a mix of new and Amazon prices is used. When available, the Amazon price is used, while when Amazon does not directly sell the product, the lowest third party price is used as indicator for the main product page price.

This approach assumes that Amazon displays its own price when possible and if no unit of the product is sold directly from Amazon, the marketplace sells automatically to the user the lowest available price among the sellers that use the Amazon marketplace to sell their products.

It is worth mentioning that Keepa does not allow users to access the full set of prices available from different sellers. Since it works as a price tracker mainly, the only price displayed from third parties is the lowest one, the one that is usually displayed in the main product page when Amazon does not directly sell the product.

3.2.3 Arellano-Bond Instruments

Due to the nature of the dataset (panel with multiple monthly lags of the independent variables) the issue of endogeneity is addressed with the use of Arellano-Bond instruments (Arellano & Bond, 1991). In this study, in fact, the variables Price, Average Rating, Number of Reviews and Amazon Choice are all endogenous and there might be a strong problem of reverse causality.[8]

To address this issue, lagged independent variables are used to instrument for prices,

[8] More on this in Section 7

average rating and number of reviews. For “Amazon Choice”, assuming that that the probability of becoming an “Amazon Choice” product depends on the popularity and customers experiences of the product, therefore “Amazon Choice” is instrumented with lagged values of average rating and number of reviews as well.

3.2.4 Rating*Reviews

Due to lack of variance in the rating variable (most average ratings are between the 4.2-4.8 stars), the interaction between average rating and number of reviews is used. This is a reliable estimate of the characteristic that a consumer looks at while viewing the Amazon search and product page. In fact both are shown side by side and the average rating is expressed as number of stars from 1 to 5 (not a precise information on the average rating) suggesting that the consumer may consider rating and number of reviews as a single characteristic.

3.3 Demographics

Demographics data used in the Random Coefficient model are taken from three different sources. There are three main demographics distributions used to simulate individuals in the European amazon markets: income, age and mobile.

Income is constructed from the EU-Silc (Eurostat, 2020) distribution of income deciles and percentiles in the European countries. From the EU-Silc, the upper bounds of each decile and of the highest 5 percentiles of income distribution are available. The agents are simulated by dividing each decile in 100 steps and every agent is assigned to a step, for the tenth decile, the upper bound is not available and therefore the 99th percentile upper bound is used as upper limit. From these 1000 individuals simulated for each market, a random draw of 100 individual incomes is taken for each market.

Income is then transformed in log of income and log of income Squared. This approach is taken from Nevo (2001), in which log(income) is supposed to have both a linear and a nonlinear impact on demand.

Age distribution is taken from the European Social Survey (*European Social Survey Round 9*, 2018) with a simple random draw of 100 individual ages for each of the five European countries

Mobile represents the prevalence of using a mobile device (phone or tablet) to shop online. The distribution of this dummy is taken from a survey conducted by the website Idealo (Pilello & De Cristofaro, 2020) on European online markets, from agents that use the price comparison tool *Idealo*[9] to shop online. From these distributions, 100 individuals are randomly drawn for each of the five European countries of interest.

All the previous variables are taken in mean demeaned form and divided by the standard error of their values across markets, this is standard in random coefficient Logit estimation and is done to normalize the coefficients of these variables.

3.4 Descriptive Statistics

The purpose of this work is to assess and quantify the effect of the "Amazon's Choice" badge on consumer's demand for products. This research was stimulated by the anecdotal evidence of the importance of the Amazon's Choice badge, as well as by the recent literature on bounded rationality and consumer choices, presented in the previous sections.

To assess this effect, a first reduced form approach can be useful in understanding whether the correlation between products popularity and products characteristics are already shown in a simple OLS regression. The purpose of this very preliminary approach is not to draw conclusions from these (clearly biased) estimates, but only to have a first look at the statistical properties of the sample used in this work.

A series of OLS regressions of Estimated Monthly Sales on products characteristics are conducted here:

$$Sales_j = -\alpha p_j + \mathbf{x}_j' \boldsymbol{\beta} + \varepsilon_j \tag{1}$$

Two approaches are proposed to assess the correlations between variables. In the first one

[9] `www.idealo.com`

(Table 1) the sample is divided in the five different markets analyzed. In the second one (Table 2), the sample is divided in five different nests of products analyzed.

Table 1: OLS Regression of Monthly Sales

	By Countries:				
	UK	Germany	Spain	France	Italy
Amazon's Choice	218.159***	164.978***	174.841***	220.540***	326.573***
	(39.760)	(38.688)	(35.131)	(34.546)	(38.156)
Bundle	50.708	95.266*	229.845***	20.277	164.088***
	(48.709)	(53.495)	(67.473)	(72.561)	(44.702)
Price	-5.703**	-0.835	-7.438***	-6.799***	-4.818**
	(2.851)	(1.944)	(2.653)	(2.177)	(1.986)
Rating*Reviews	0.030***	0.016***	0.000	-0.001	0.002*
	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)
Constant	202.000***	212.949***	293.943***	278.024***	223.992***
	(43.394)	(40.151)	(36.685)	(36.498)	(36.065)
Observations	409	361	355	362	427
R-squared	0.203	0.095	0.123	0.129	0.236

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The results show a persistent large positive correlation between Amazon's Choice badge and quantities of products sold in a month. The impact tends to fluctuate a lot between countries as well as between different groups of products. This approach does not allow to draw any causal relationship between the variables, it however shows that Amazon's Choice products are consistently more popular than the rest of the products. This was the expected result indeed and is useful in testing the quality of the sample proposed.

4 Theoretical Model

The theoretical model presented is a standard discrete choice model, in which agents choose the product which gives them the highest utility, accounting for agent specific variations in the coefficients of the products characteristics.

Table 2: OLS Regression of Montly Sales

	By Nests:				
	Dvi	Ethernet	Hdmi	Lightning	Usb
Amazon's Choice	-27.473	11.546	21.226	176.818***	367.718***
	(31.463)	(16.644)	(24.771)	(47.407)	(27.045)
Bundle	-15.871	-2.177	27.663	60.478	161.267***
	(81.261)	(23.609)	(44.749)	(62.052)	(34.794)
Price	4.612	-0.820	1.316	-5.289	-5.247***
	(2.841)	(0.964)	(1.510)	(4.394)	(1.411)
Rating*Reviews	0.055***	0.006**	0.000	0.001	0.016***
	(0.014)	(0.003)	(0.001)	(0.002)	(0.003)
Constant	122.603***	223.991***	239.241***	485.077***	150.840***
	(37.967)	(20.074)	(24.312)	(61.591)	(24.317)
Observations	61	308	324	366	855
R-squared	0.244	0.019	0.007	0.043	0.305

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Following the literature that this work builds on, first a (behavioral) specification of the discrete choice model is presented, in which consumers are divided in two types, with one type only considering the Amazon's Choice product and ignoring all others. This specification builds onto the Behavioral Industrial Organization tradition. Then, the model is further developed in a fully rational specification, with consumers considering all products and choosing which one to buy according to a standard utility maximization problem. This specification builds on the literature of the discrete choice models in Empirical Industrial Organization applications.

The model starts by considering a single agent i who chooses among a set of j products, according to a utility function which is linear in product characteristics:

$$\mu_{i,j} = \alpha_i(y_i - p_j) + \mathbf{x}_j'\boldsymbol{\beta}_i \tag{2}$$

Where y_i is consumer i's income, p_j is product j price and $\mathbf{x}_j$ the vector of product j characteristics, with dimensions Kx1. The α_i and $\boldsymbol{\beta}_i$ coefficients are individual specific. To simplify the analysis and to make the model tractable in the empirical part, the average utility is separated from the individual specific portion of utility.

$$\mu_{i,j} = \alpha(y_i - p_j) + \mathbf{x}_j'\boldsymbol{\beta} + \epsilon_{i,j} \tag{3}$$

Where $\boldsymbol{\beta}$ is the average $\boldsymbol{\beta}_i$ coefficient. and α is the average α_i coefficient. So that the average utility of product j is $\alpha(y_i - p_j) + \mathbf{x}_j'\boldsymbol{\beta}$, while $\epsilon_{i,j}$ is the individual deviation from the mean utility of product j.

Each user only buys one product, the one that maximize its utility function. Before looking at all products, the user is immediately redirected to the Amazon's Choice offer, either automatically (by voice assistants such as Alexa) or by looking directly at the page, in which the Amazon's Choice product immediately stands out among the others.

There are two ways to model how consumers react to the Amazon's Choice badge: the more modern bounded rationality approach and the traditional full rationality approach.

4.1 Bounded Rationality

The bounded rationality model presented here is a variation of Stahl (1989) model of competition with sequential consumer search. In Stahl's model, consumers compare different products sequentially and firms compete in price. Consumers are divided in two types: the first type of consumer does not bear any search costs and looks for the product offered at the lowest possible price. The second type of customer bears search costs and, as proven by Stahl, chooses the first product with a price lower than his "reservation price". The reservation price is set as a function of the search costs.

Stahl's model would not be appropriate in this setting, as there are virtually no search costs in the Amazon's marketplace and all results are displayed at the same time.[10] Instead, the model proposed here takes Stahl's intuition and adapts it to a setting in which a portion

[10] Assuming that the costs of switching between results pages is close to zero.

of consumers prefers to consider only the Amazon's Choice option and buys it if the expected utility is larger than a minimum (reservation) level.

In this setting, there are two types of customers: one tends to buy the Amazon Choice in any case and the other looks at all other products, without being influenced by the Amazon Choice badge. Each consumer is assigned randomly to group A ("Amazon Choice") or group B ("Rational Consumer"). In the first group, the user buys the Amazon Choice product if the utility of the product is more than a minimum level of utility set at $\mu_{i,0}$:

$$
\begin{aligned}
\max_{j\in\{a,0\}} \quad & U_{i,j} = \mu_{i,j} - \mu_{i,0} \\
\text{sub:} \quad & \mu_{i,a} = \alpha(y_i - p_a) + \mathbf{x}_a'\boldsymbol{\beta} + \epsilon_{i,a} \\
& \mu_{i,0} = \mu_0 + \epsilon_{i,0} \\
& \mu_0 \geq 0
\end{aligned}
\tag{4}
$$

The choice of product j for agent i is then:

$$
j_i = \begin{cases} a & \mu_{i,a} \geq \mu_{i,0} \\ 0 & \mu_{i,a} < \mu_{i,0} \end{cases} \tag{5}
$$

Agent B instead chooses the product with the greatest utility according to a traditional maximization function:

$$
\begin{aligned}
\max_{j\in\{0,1,\dots J\}} \quad & U_{i,j} = \mu_{i,j} - \mu_{i,0} \\
\text{sub:} \quad & \mu_{i,j} = \alpha(y_i - p_j) + \mathbf{x}_j'\boldsymbol{\beta} + \epsilon_{i,j} \qquad \forall j \in \{1, 2, \dots J\} \\
& \mu_{i,0} = \mu_0 + \epsilon_{i,0} \\
& \mu_0 \geq 0
\end{aligned}
\tag{6}
$$

4.2 Full Rationality

The alternative setting is the traditional fully rational consumers setting, in which all customers look at the utilities of the products and choose according to a traditional maximization function. Here "Amazon Choice" is considered a characteristic with a linear impact on the demand of product j:

$$\begin{aligned} \max_{j\in\{0,1,...J\}} \quad & U_{i,j} = \mu_{i,j} - \mu_{i,0} \\ \text{sub:} \quad & \mu_{i,j} = \alpha(y_i - p_j) + \mathbf{x}_j'\boldsymbol{\beta} + A_j\beta_a + \epsilon_{i,j} \qquad \forall j \in \{1, 2, ...J\} \\ & \mu_{i,0} = \mu_0 + \epsilon_{i,0} \\ & \mu_0 \geq 0 \end{aligned} \tag{7}$$

Where A_j is the dummy that indicates whether the product is "Amazon Choice" or not. Consumer heterogeneity is accounted for by $\epsilon_{i,j}$.

In the next section, the coefficients of the utility functions of the consumers are estimated. All utilities functions previously displayed can be simplified for a cleaner notation.

$$\begin{aligned} U_{i,j} &= \mu_{i,j} - \mu_{i,0} \\ &= \alpha(y_i - p_j) + \mathbf{x}_j'\boldsymbol{\beta} + A_j\beta_a + \epsilon_{i,j} - \mu_0 - \epsilon_{i,0} \end{aligned} \tag{8}$$

μ_0 is accounted for in the constant term of our econometric equation and a single ϵ error term is considered for simplicity, so the variables μ_0 and $\epsilon_{i,0}$ are removed. Finally, the utility of the agent buying product j is simply:

$$U_{i,j} = \mu_{i,j} = \alpha(y_i - p_j) + \mathbf{x}_j'\boldsymbol{\beta} + \epsilon_{i,j} \tag{9}$$

Where $\mathbf{x}_j$ includes the A_j "Amazon Choice" dummy. and the utility of consumer buying the outside good can be simply written as:

$$U_{i,0} = \mu_{i,0} = 0 \tag{10}$$

5 Empirical Approach

The empirical approach for the paper follows the work of S. T. Berry (1994), S. Berry, Levinsohn, & Pakes (1995), and the more recent works on the matter of Nevo (2001) and Rasmusen et al. (2007).

Two different demand equations are estimated. First the behavioral consumer, as shown in models (4) and (6): Here, the "Amazon Choice" characteristic is not part of the characteristics set (K), but is used instead as a determinant for the nest division in the nested logit approach.

$$\mu_{i,j} = \alpha(y_i - p_j) + \mathbf{x}_j'\boldsymbol{\beta} + \epsilon_{i,j} \tag{11}$$

For the second one instead, where all consumers are considered to be perfectly rational, the "Amazon Choice" dummy A_j is part of the characteristics set.

$$\mu_{i,j} = \alpha(y_i - p_j) + \mathbf{x}_j'\boldsymbol{\beta} + \beta_a A_j + \epsilon_{i,j} \tag{12}$$

To estimate the first model, this work resorts to a three levels nested logit, while to estimate the second one, first a more traditional two levels nested logit and then a more complex Random Coefficient approach are used. (S. Berry, Levinsohn, & Pakes, 1995).

The models are estimated in parallel to assess whether the estimates of the impact of the Amazon's Choice badge on consumers' choices are robust to changes in the theoretical model of consumers' choice.

5.1 Multinomial Logit

The simple multinomial logit approach is the starting point for all the subsequent models. It starts with the utility of agent i buying product j from equation (11).[11]

Assuming agents are fully rational (this assumption will be relaxed later in the three level nested logit[12]), agent i will buy product j if and only if $\mu_{i,j} \geq \mu_{i,k} \forall k \neq j$. This allows to simplify equation (11) by noting that y_i is eliminated in considering the optimal choice for agent i. The term ε_j is included and represents the mean deviation of agent i from the predicted utility of good j, so then $\epsilon_{i,j}$ now represents the individual deviation from the mean deviation from the predicted utility.

$$\mu_{i,j} = \mathbf{x}_j' \boldsymbol{\beta} - \alpha p_j + \epsilon_{i,j} + \varepsilon_j \tag{13}$$

Assuming each agent will buy only one quantity of product j, the quantity of products j sold is expressed as:

$$q_j = M * p_j(\mathbf{x}, \boldsymbol{\varepsilon}, \mathbf{p}, \theta) \tag{14}$$

By dividing both sides for the market size M, the market share of product j (d_j) is:

$$d_j = q_j / M = p_j(\mathbf{x}, \boldsymbol{\varepsilon}, \mathbf{p}, \theta) \tag{15}$$

As shown by Cardell (1997), through the simple assumption of $\epsilon_{i,j}$ IID and distributed as EV type 1, it is possible to express the predicted market shares (d_j) in the Logit functional form:

$$d_j = \frac{exp\{\delta_j\}}{\sum\limits_{j=0}^{J} exp\{\delta_j\}} \tag{16}$$

Where δ_j is the mean utility of product j: $\delta_j = \mathbf{x}_j' \boldsymbol{\beta} - \alpha p_j + \varepsilon_j$ and δ_0 is the mean utility of

[11]For simplicity we only refer to equation (11) but this model can easily be extended to (12) simply by including A_j in the characteristics set K

[12]see section 5.2.1

the outside good.

By assuming $\delta_0 = 0$:

$$\begin{aligned} d_j &= \frac{exp\{\delta_j\}}{1 + \sum_{j=1}^{J} exp\{\delta_j\}} \qquad \forall j \in 1, 2, ...J \\ d_0 &= \frac{1}{1 + \sum_{j=1}^{J} exp\{\delta_j\}} \end{aligned} \tag{17}$$

A closed form solution for the mean utility δ_i is easily found by taking the difference in logs of d_j and d_0.

$$\ln(d_j) - \ln(d_0) = \mathbf{x}_j' \boldsymbol{\beta} - \alpha p_j + \varepsilon_j \tag{18}$$

Substituting the predicted market shares d with the observed market shares S from the data, the econometric equation of the logit model is found:

$$\ln(S_j) - \ln(S_0) = \mathbf{x}_j' \boldsymbol{\beta} - \alpha p_j + \varepsilon_j \tag{19}$$

5.1.1 IIA

Unfortunately the above equation, while being easy to implement and useful for some initial estimates, cannot is not particularly robust due to a very restrictive consequence of IID EV1 errors: Independence of Irrelevant Alternatives. In fact it can be proved that the result of estimating equation (19) the ratio of two products market shares have this property (Train, 2009):

$$\frac{d_j}{d_k} = \frac{e^{\delta_j} / \sum_j e^{\delta_j}}{e^{\delta_k} / \sum_j e^{\delta_j}} = \frac{e^{\delta_j}}{e^{\delta_k}} \tag{20}$$

The result is that when a characteristic of product k changes (say price), all other products shares react in the same manner:

$$\frac{\partial q_k}{\partial p_j} = -\alpha p_k d_k \quad \forall j \neq k \tag{21}$$

This is unrealistic in the setting of this work, as various groups of products with different purposes are considered here. The demand for all these products is not likely to react in a homogeneous way to a change in characteristics of single product. To mitigate this issue, it is necessary to turn to the more advanced approaches proposed below.

5.2 Nested Logit

In the nested Logit approach, the products are divided in different groups ("Nests"), and the choice of consumers is modelled in the sequential manner that follows:

1. Consumers choose which Nest of products to buy.

2. Inside each nest, consumers choose by maximizing their utility function.[13]

This approach requires the researcher to know the different categories in which the products are divided according to the customers. While in some cases it can be a limitation, in this setting the Amazon marketplace already divides its products in reliable subcategories. For this work, the 5 subcategories of usb cables sold on Amazon with the most products in them (usb, ethernet, dvi, lightning, hdmi) are used as the five nests of the analysis.

The separation of different products into nests is fundamental to model with more precision the consumer heterogeneity and to restrict the span of the IIA restrictions posed earlier. IIA property in fact now holds only within the single nests and not across the different nests (Train, 2009).

The utility of agent i can now be expressed as:

$$\mu_{i,j} = \mathbf{x}_j'\boldsymbol{\beta} - \alpha p_j + \zeta_g + (1-\sigma)\epsilon_{i,j} + \varepsilon_j \tag{22}$$

Where ζ_g is the mean utility of all products belonging to the same group (or nest) $g \in G$. The probability that agent i buys product j now needs to be expressed in a sequential manner

[13] Note: this model still uses the full rationality approach, for the behavioral model see the next section

that reflects the nested form:

$$d_j = d_g * d_{j|g} \tag{23}$$

According to S. T. Berry (1994), by assuming that $\zeta_g + (1-\delta)\epsilon_{i,j}$ follows an IID, EV1 distribution, d_g (the average probability that agents chooses nest g) and $d_{j|g}$ (the average probability that agents choose product j inside nest g) can be expressed as:

$$\begin{aligned} d_{j|g} &= \frac{exp\{\delta_j/1-\sigma\}}{D_g} \\ d_g &= \frac{D_g^{1-\sigma}}{\sum\limits_{g\in G} D_g^{1-\sigma}} \end{aligned} \tag{24}$$

Where:

$$D_g = \sum_{j\in g} exp\{\delta_j/(1-\sigma)\} \tag{25}$$

By expressing the outside good as a separate group, the predicted market share of the outside good is:

$$d_0 = \frac{1}{\sum\limits_{g\in G} D_g^{1-\sigma}} \tag{26}$$

From (23) it is easy to derive that:

$$\ln(d_j) - \ln(d_0) = \frac{\delta_j}{1-\sigma} - \sigma \ln(D_g) \tag{27}$$

By adding $\sigma d_{j|g}$ and substituting the predicted with the observed market shares:

$$\ln(S_j) - \ln(S_0) = \sigma \ln(S_{j|g}) + \mathbf{x}_j'\boldsymbol{\beta} - \alpha p_j + \varepsilon_j \tag{28}$$

Where the coefficient σ is restricted between 0 and 1 and represents the relative importance of the division in nests against the products characteristics. The closer σ is to 1, the more it is important for customers the division in nests; the more it is close to 0, the more the individual products characteristics are important for consumers. With the limit case $\sigma = 0$, the two levels nested logit equation (28) becomes identical to the simple logit equation (19).

5.2.1 Three Levels Nested Logit

To model the behavioral type of consumer, a three level nested logit is used. Agents now choose according to the following order:

1. Consumers choose which Nest of products to buy.
2. Behavioral Consumers choose the Amazon's Choice suggestion, while the Rational Consumers choose among the other products.
3. Consumers compare all alternatives in each subnest and buy the option which maximizes their utility function.

The utility function of consumer i will therefore look like this

$$\mu_{i,j} = \mathbf{x}_j'\boldsymbol{\beta} - \alpha p_j + \varepsilon_j + \epsilon_{i,j} \tag{29}$$

where $\epsilon_{i,j}$ accounts for both the nest and the amazon choice fixed effects. Under the usual IID EV1 assumption, the closed form solution for the econometric equation is then:

$$\ln(S_j) - \ln(S_0) = \mathbf{x}_j'\boldsymbol{\beta} - \alpha p_j + \sigma_{hg}\ln(S_{j|h}) - \sigma_g \ln(S_{h|g}) + \varepsilon_j \tag{30}$$

The proof for the closed form solution of the three level nested logit is derived in the Appendix.

A potential criticism to this approach is that here the full choice set of the rational consumer is not represented. In fact the rational consumer can only buy the non Amazon's Choice products with the three level nested Logit. The ideal approach would be to use a Cross Nested Logit (Ben-Akiva & Bierlaire, 1999). With the Nested Logit approach, in fact the model only allows products to be included in a single sub-nest. With that restriction, Amazon Choice products can only be considered by the consumers blindly following the Amazon default choice, while other consumers which evaluate products according to a traditional utility maximization problem are restricted in never looking at Amazon's Choice products. With the cross nested Logit approach the Amazon's Choice products could be included both in the rational consumer's choice set and in the behavioral consumer's choice set. To do so, however, it would be necessary to have the correct proportion for the share of the second level subnests: $S_{h|g}$. This would means that in the cross nested Logit approach, the proportion of behavioral customers to rational customers would be needed for every nest. This proportion is likely impossible to derive without micro data on the behavior of customers on the Amazon platform. The proposed three level nested Logit, instead, while imposing an additional stringent assumption, is more tractable, as the ratio of behavioral customers to rational customers is by construction equal to the ratio of market shares of Amazon's Choice to non Amazon's Choice products.

5.3 Random Coefficient Logit

The Random Coefficient Logit model allows to specify more precisely the individual specific portion of the Utility, resulting in more accurate estimates of the mean utility function δ_j, together with a second set of coefficients of products characteristics, interacted with individual random and demographic variables. First introduced by S. Berry, Levinsohn, & Pakes (1995), it was then made a popular option for demand estimation by Nevo (2001). Since then, the Random Coefficient Logit has had various applications, especially in the field of applied Industrial Organization.Decarolis (2015) for the application side and Rasmusen et al. (2007) on the methodological part are also extremely important and are followed for the setup of this work.

This setting starts once again with the usual individual utility equation, but the coefficients α_i and β_i are individual specific this time, and the notation t indicates market $t \in T$.

$$\mu_{i,j,t} = \mathbf{x}'_{j,t}\boldsymbol{\beta}_i - \alpha_i p_{j,t} + \epsilon_{i,j} + \varepsilon_j \tag{31}$$

Instead of aggregating the individual specific portion of the demand in the variable $\epsilon_{i.j}$, or specifying group specific utilities as in the nested logit, now the individual portion of the coefficients is considered explicitly:

$$\begin{pmatrix} \alpha_i \\ \boldsymbol{\beta}_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \boldsymbol{\beta} \end{pmatrix} + \begin{pmatrix} \sum_\alpha \\ \sum_\beta \end{pmatrix} \boldsymbol{v}_i + \begin{pmatrix} \prod_\alpha \\ \prod_{\boldsymbol{\beta}} \end{pmatrix} \boldsymbol{D}_i \tag{32}$$

Where v is a kx1 vector of normal and randomly distributed variables and D_i is the mx1 vector of demographic variables for agent i. $\sum$ is the kxk matrix of interactions between random variables v and products characteristics X, and $\prod$ is kxm matrix of interactions between demographic variablels D and products characteristics X. The inclusion of demographic variables was proposed in Nevo (2001) and has since become the standard for Random Coefficient Logit estimation.

From 32 it is possible to derive the following utility equation:

$$\mu_{i,j,t} = \delta_{j,t} + u_{i,j,t} + \epsilon_{i,j,t} \tag{33}$$

Where:

$$\begin{aligned} \delta_{j,t} &= \mathbf{x}'_{j,t}\boldsymbol{\beta} - \alpha p_{j,t} + \epsilon_{j,t} \\ u_{i,j,t} &= (-p_{j,t} \quad \mathbf{x}'_{j,t}) \sum \boldsymbol{v}_i + (-p_{j,t} \quad \mathbf{x}'_{j,t}) \prod \boldsymbol{D}_i \end{aligned} \tag{34}$$

And $\varepsilon_{i,j,t}$ is still assumed to IID with EV Type 1 distribution. This setting offers no closed form solution to extract the mean utility, like it was possible to do in the Logit approach. To estimate equation (33) it is therefore necessary to procede with an optimization method explained here below:

Before starting with the optimization algorithm, it is necessary to specify the starting values of the coefficients α and $\boldsymbol{\beta}$ in $\delta_{j,t}$ and for the coefficient $\sum$ and $\prod$ in $u_{i,j,t}$. In this work, α and $\boldsymbol{\beta}$ starting points are derived by applying the simlple logit regression (19). What follows is the step by step explaination of the optimization routine used to calculate the Random Coefficient logit coefficients estimates, it is taken mostly from Nevo (2001) and Rasmusen et al. (2007):

0. Start by making random draws of the v_i's and D_i's. The distribution of v_i are assumed standard normal, while the distributions of D_i's are taken from the real distributions of the demographics in the different markets. This step will only be done at the beginning of the estimation and will not be part of the optimization routine.

1. Calculate the predicted market shares:

$$d_{jt} = \frac{1}{n} \sum_{i=1}^{n} \frac{exp\{\delta_{j,t} + u_{i,j,t}\}}{1 + \sum_{j=1}^{J} exp\{\delta_{j,t} + u_{i,j,t}\}} \tag{35}$$

2. Use the following contraction mapping to get a better estimate of δ

$$\boldsymbol{\delta}^{h+1} = \boldsymbol{\delta}^{h} + (\ln(\mathbf{S}_t) - \ln(\mathbf{d_t})) \tag{36}$$

   Where $\mathbf{S}_t$ are the real market shares and $\mathbf{d}_t$ are the predicted market shares in market t calculated in (1). When the contraction mapping is concluded, the result is a more precise estimate for the mean utility $\boldsymbol{\delta}_t$, called $\hat{\boldsymbol{\delta}}_t$

3. Calculate the error term

$$w_{j,t} = \hat{\delta}_{j,t} - (-\alpha p_{j,t} + \mathbf{x}'_{j,t}\boldsymbol{\beta}) \tag{37}$$

   And obtain the moment condition:[14] $\boldsymbol{w'Z\Omega^{-1}Z'w}$.

4. Compute better estimates for $\alpha, \beta, \sum, \Omega$.

   - Find $(\hat{\alpha}, \hat{\beta})$ such that:

$$(\hat{\alpha}\ \hat{\boldsymbol{\beta}})_{GMM} = (\boldsymbol{X'Z\Omega^{-1}Z'X})^{-1}\boldsymbol{X'Z\Omega^{-1}Z'\delta} \tag{38}$$

[14]Z is the matrix of exogenous variables, used as instruments for price and for the other endogenous characteristics in X (if present)

- With the new estimates $\hat{\alpha}$ and $\hat{\beta}$ get the new error term:

$$\hat{w}_{j,t} = \hat{\delta}_{j,t} - (-\hat{\alpha} p_{j,t} + \mathbf{x}'_{j,t} \hat{\boldsymbol{\beta}}) \tag{39}$$

- Get the new moment condition $\boldsymbol{\hat{w}' Z \Omega^{-1} Z' \hat{w}}$ with the new $\hat{\alpha}$ and $\hat{\boldsymbol{\beta}}$ just calculated
- Estimate the more precise Hansen (1982) weighting matrix with the new error term $\boldsymbol{w}$: $\boldsymbol{\Omega = (E(Z'ww'Z))}$.
- Use a search algorithm, iterating from 1 to 4, to find the best estimate for $\sum$ amd $\prod$, which is the one that minimizes the GMM objective function[15].

$$\hat{\boldsymbol{\theta}} = \arg\min_{\theta} \boldsymbol{w' Z \Omega^{-1} Z' w} \tag{40}$$

Note that $(\hat{\alpha}, \hat{\beta})$ are calculated linearly at (38), while $\sum$ and $\prod$ are obtained through the search algorithm by iterating from point 1 to 4.

6 Identification Strategy

6.1 The Endogenous Prices Problem

The problem of identification in a demand estimation setting rises immediately due to the nature of competition in free markets. Prices are in fact never exogenous, unless they are predetermined from a regulatory source, or they are randomly assigned. In all markets in which there is an interplay between demand and supply, high prices could be caused by high demand and low prices could be caused by low demand. To tackle this problem, the ideal setting would be to use an exogenous shock to prices as an instrument for a two-stage-least-squares or a GMM approach. In the proposed setting, this is not possible as no price shocks

[15] This work follow Nevo (2001) suggestion of using a simplex search, starting from guessed $\sum$ and $\prod$ values, then use the more precise (but more sensibe to the starting point) Quasi-Newton optimization algorithm to estimate $\sum$ and $\prod$, using as a starting point the $\sum$ and $\prod$ coefficients obtained with the simplex search method

are available in the time-frame analyzed. Another, equally valuable, approach would be to use production costs as instruments for prices. Such approach would be extremely valuable, but the lack of availability of cost data for large products datasets renders it a relatively unused option.

In the applied industrial organization literature, three different approaches are systematically used to solve the endogenous prices problem. Three types of instruments can be constructed according to the specific characteristics of the used datasets: Arellano-Bond instruments (Arellano & Bond, 1991), BLP instruments (S. Berry, Levinsohn, & Pakes, 1995) and Haussman-Nevo (Nevo, 2001) instruments. Arellano-Bond instruments exploit time changes of endogenous characteristics when a dynamic panel data of products characteristics is available. BLP instruments are created from a set of neighbouring products characteristics. Haussman-Nevo instruments are derived from the regional variations in prices.

This paper exploits the availability past prices to instrument the current price levels. This approach is therefore a variation of what was proposed in Arellano & Bond (1991). Prices in every period are determined by the costs of production of the product (c), plus a reaction to the current demand (δ_t): when demand is high prices tend to increase, when demand is low prices tend to decrease. Since, at least in the short term, costs can be considered fixed, then past prices and current prices are highly correlated.

$$p_t = c + \gamma\delta_t + \varepsilon_t \tag{41}$$

This leads to past prices being a strong instrument for current prices.

It is however necessary to verify that past prices are exogenous and affect current demand only through the current prices. This last assumption (better known as the "Exclusion Restriction") is easy to verify as customers only see current prices and are not considering past prices when making a purchase decision.[16] Therefore, the Exclusion Restriction is satisfied by

[16] A case can be made that explicit discounts that show both past prices and current prices would indeed render the assumption false as the customer could be influenced by the amount of discount applied to the product, transforming the average portion of the demand equation in: $\delta_t = -\alpha_1 p_t + -\alpha_2(p_t - p_{t-1}) + \mathbf{x}_t'\boldsymbol{\beta} + \varepsilon_t$.

construction of our model and there is no need to assume that past prices can be considered as a relevant characteristic, when customers decide to buy. Since past demand is correlated with current demand, the exclusion restriction could be invalidated if low prices lead to high demand in the past and then current demand stays high regardless of current prices. The only channel through which this could be possible is through reviews and Amazon-Choice selection: as products become more popular they receive better reviews and therefore the demand stays high even when prices go up again. Or something similar could happen to a product which becomes popular due to low prices and then is given the Amazon's Choice badge, once the prices go up demand could stay high anyways due to the badge. These cases are treated by including all time varying product page characteristics in the econometric equation. By including current characteristics in the second stage of the instrumental variable approach, past prices can only affect current demand through current prices and current characteristics.

The last, and most important, caveat of this approach is the need to assume that errors in the demand function are not serially correlated. The main problem that drives the endogeneity of prices is reverse causality, prices, however, could also be correlated with omitted variables in the demand equation. it is therefore necessary to assume that the error term ε_j in the average demand equation is not serially correlated:

$$\delta_j = \mathbf{x}_j' \boldsymbol{\beta} - \alpha p_j + \varepsilon_j \tag{42}$$

While writing this work, no reason for serially correlated errors was found, nor any reason to think that any omitted variable, correlated with price of characteristics, could be affecting current demand. However, as an additional precaution, first monthly lags were not used and instead to instrument for current prices, second monthly lags of prices were used. By using preceding months as instruments, no relevant differences were found in the estimation results.

In this work however agents are considered to be mostly rational in their utility function and a discount applied to the product does not induce direct utility gains on a rational consumer

6.2 Extending the Framework to All Endogenous Characteristics

In this framework, prices are not the only endogenous variable, unlike the standard applied Industrial Organization problems. In fact, in the Amazon online market, agents are exposed to other three endogenous characteristics: number of reviews, average rating of the reviews and Amazon's Choice badge.

The first two are interacted to construct the rating*reviews coefficient. The reason for this choice is that the rating is not particularly explicit in the Amazon Marketplace (being indicated only as stars) and overall not a high variance in ratings was found. Since agents see average rating and number of reviews one next to the other it can be assumed that they consider rating and reviews only in conjunction one with the other (a single 5 stars review is probably worth less than one hundred reviews with 4.5 stars on average). Rating*reviews is still endogenous and is therefore instrumented with a number of past lags of itself, starting from the second lag.

Amazon's Choice is the most difficult to instrument as there is no clear indication from Amazon on how it is assigned. For that reason, the most plausible reason is followed here, more popular and better performing products are more likely to be assigned the Amazon's Choice badge than less popular and worse rated products. For that reason, Amazon's Choice is instrumented with the same lags of rating*reviews explained in the previous paragraph. More than one lag is included for this reason as well: at least one lag is necessary to instrument rating*reviews, at least another lag is necessary to instrument Amazon's Choice.

All the previous considerations on the validity of the price instruments hold here as well for the identification of the characteristics coefficients

6.3 Brand Fixed Effects and Bundle

Brand fixed effects are included to account for products produced by the same popular brand. Only the 20 most popular brands are included to create a dummy matrix, while all other brands are considered the additional "Other" brand. The reasons for this approach are both theoretical and operational. On the operational side: it is preferable to avoid including too many firm dummies to avoid to drastically decrease the degrees of freedom and inflate the coefficients variances. On the theoretical side, products coming from less popular brands are most likely considered by customers as relatively similar, while only popular brands may bring an appreciable " brand reputation" effects on the utility of consumers.

The bundle characteristic is included by scraping the title of products to find products that are offered in a set of more than one unit. It is an extremely important variable as it is not uncommon that these types of electronic disposable products are offered in such form. The utility of a set is likely less than two times the utility of the unit, as the marginal utility of a second unit of the same product is likely much lower than the marginal utility of the first product.[17] Nevertheless, this additional utility must be taken into account and for this reason the bundle dummy is added to the characteristics set.

7 Results

This analysis led to interesting and sometimes unexpected results on the demand estimation of Amazon products, what follows is an overview and comment on the main results of the model.

All the following models include the same characteristics X and the same set of instruments Z. The dataset includes the 2500 products which come from the five different Amazon EU markets (Italy, France, Spain, Germany, United Kingdom). From these five markets, the 100 most sold products in each of these electronics cables subcategories are taken (usb, hdmi, vga, lightning, ethernet).

USB cables are considered the outside good, this is necessary to find a simple closed form solution for our nested Logit. The standard category of "usb" cables was chosen because it contains the most standardized products and is very likely that an agent buying in the

[17]The marginal utility of the first unit derives from having a product that the agent needs, while the marginal utility of the second unit is the one of storing a product which the agent may be needing in the future (but the agent doesn't know yet if the second stored unit will be ever used)

"usb" nest is looking for any product that could fulfill his needs, without looking at the differentiation among products. The utility of the nest "usb" is then standardized to zero, while the mean utility of all other products is estimated.

All products with a market share of less than 0.05% are then added to outside good and only products for which at least the last five months of data is available from the keepa api. After this and excluding the outside goods, the sample used contains 1059 products.

7.1 Multinomial Logit

The first technique implemented is the simple Multinomial Logit, thanks to its easy to implement closed form solution, it provides a valuable starting point for the analysis. Here is reported the functional form of the econometric equation for completeness:

$$\ln(S_{j,t}) - \ln(S_{j,0}) = -\alpha_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \varepsilon_j \tag{43}$$

Before using the generalized method of moments approach to instrument for endogenous variables, it is important to start by analyzing the impact of products characteristics on consumers' utilities in a simple OLS regression. Note that from an econometric point of view this regression is not correct, as there are multiple variables in X which are endogenous to the demand, nevertheless, the OLS regression here is included to assess the correlations between the products characteristics and the utility of agents. In model (1) of Table 7.1 it is possible to see that there is a negative correlation between price and Utility, while Amazon Choice, Bundle and Rating*Reviews have all a positive effect. The effect of Amazon Choice here seems already quite large as it can be comparable to offering a product in a bundle. The correlation between price and Utility is instead quite small, same applies to the correlation between the Rating*Reviews of the products and the demand. It is important to remember, however, that apart from the bundle coefficient, all others are likely to be biased in this first stage. Price may suffer a positive bias due to the endogeneity with demand

Table 3: Multinomial Logit Estimates

	(1) OLS	(2) GMM	(3) GMM
Amazon's Choice	0.160***	0.0516	1.098*
	(0.0446)	(0.545)	(0.657)
Price	-0.00424	-0.00465	0.000541
	(0.00308)	(0.00321)	(0.00730)
Rating*Reviews	2.07e-06	3.72e-06	-2.41e-06
	(1.34e-06)	(2.46e-06)	(3.58e-06)
Bundle	0.117*	0.135	0.116
	(0.0667)	(0.0870)	(0.104)
Constant	-5.495***	-5.442***	-6.135***
	(0.0500)	(0.289)	(0.471)
Observations	1,059	1,059	1,059
R-squared	0.021	0.015	
Hansen's J		1.922	0
Hansen p_value		0.166	1
Brand FE			Yes

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

more popular products may be more expensive due to the high demand. The same can be said for the Rating*Reviews coefficient: products that were more popular in the past may receive more reviews, therefore showing a positive correlation between Rating*Reviews and market shares, which translates in a positive correlation between Rating*Reviews and utility. Amazon Choice could be biased on one side due to reverse causality with the demand: high past demand make a product more likely to be selected as Amazon Choice, therefore Amazon Choice will show high correlation with the current demand. On the other side, however, it is not possible at this stage to know if there are other factors influencing the assignment of the Amazon Choice badge, some products may be excluded from the selection due to low availability in the region and difficulties in shipping from Amazon for example.

For all these reasons it is important to implement a different approach, the equations that follow are in fact estimated with a standard Generalized Method of Moments procedure, in which the independent variables are instrumented with the two months lag of Price and the two months, three months and four months lags of Rating*Reviews. The reason for this is to include one instrument for Price, one instrument for Rating*Reviews, at least one instrument for the Amazon Choice badge and the extra lags of the Rating*Reviews variables are included in order to allow for the Hansen and Sargan test for over identifying restrictions. (Sargan, 1958)

In order for these estimates to be unbiased, it is necessary to assume that the error terms of equation (43) are not serially correlated for more than two months. Under this restriction, then the three assumptions of the GMM should all be satisfied.[18]

As shown in model (2) of table 7.1, without the brand fixed effects none of the coefficients is statistically significant, and they are all positive and small, apart from price which remains negative. Including brand fixed effects in model (3), Amazon Choice seems to have a much larger impact than before and is statistically significant at the 10% level. However none of the other relevant variables is statistically significant and price even displays a positive coefficient.

These estimates may be unreliable also due to the IIA property[19] which is unrealistic given that in this setting the sample includes electronic cables made for different purposes. In fact the IIA property leads to the conclusion that a change in a single characteristic of one of the products will lead to the same change in quantities sold of all other products. For example this leads to assume that if an “hdmi” cable increases in price, then both a direct competitor “hdmi” cable and a “lightning” cable will gain the same quantities sold. This is not possible in this setting as these cables have specific different purposes, even if they are under the same market of products. It is necessary therefore to turn to a more robust method (the Nested Logit) which should allow for a better model of consumer heterogeneity.

[18]See section 6.1

[19]See Section 5.1.1

7.2 Nested Logit

The Nested Logit approach allows to better model the heterogeneity of individual consumers and it should give more precise estimates for the demand coefficients. Note here that the IIA property holds only between products of the same group and therefore is a more reasonable assumption given our dataset of highly substitutable products inside each category. This section starts with the two-levels Nested Logit, under which it is assumed that all agents are rational utility maximizers. Later a third level is included, in which agents are either "Behavioral" or "Rational", the "Behavioral" agents will only consider the Amazon Choice good, while the "Rational" agents will consider all other options.

Table 4: Nested Logit Estimates

	(1) Two Levels	(2) Two Levels	(3) Three Levels	(4) Three Levels
Amazon's Choice	-0.0224	1.174**		
	(0.440)	(0.465)		
Price	-0.00449*	-0.000378	-0.00442*	-0.00904***
	(0.00242)	(0.00564)	(0.00240)	(0.00272)
Rating*Reviews	4.95e-06**	-6.99e-07	4.88e-06***	3.57e-06**
	(1.99e-06)	(2.47e-06)	(1.29e-06)	(1.47e-06)
Bundle	0.222***	0.220**	0.224***	0.271***
	(0.0703)	(0.0910)	(0.0565)	(0.0637)
σ	0.493***	0.631***		
	(0.0299)	(0.0496)		
σ_g			0.413***	0.507***
			(0.0607)	(0.0599)
$\sigma_{h,g}$			0.498***	0.590***
			(0.0302)	(0.0351)
Constant	-3.258***	-3.396***	-3.309***	-2.883***
	(0.271)	(0.289)	(0.149)	(0.150)
Observations	1,059	1,059	1,059	1,059
R-squared	0.291		0.296	
Hansen's J	4.218	0	4.098	4.750
Hansen p_value	0.0400	1	0.129	0.0930
Brand FE		Yes		Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

7.2.1 Two Levels

The results of the two-levels Nested Logit are reported in models (1) and (2) of table 7.2. It is easy to see that the coefficients are for the majority larger than before and statistically significant. One important different is that in model (2) the impact on consumers' utility of the Amazon's Choice badge and of the bundle is much larger than before. Price looses statistical significance in model (2) as well as Rating*Reviews. But nevertheless the signs of the coefficients are consistent with what was shown in the Multinomial Logit. Note that in model (1), which does not include the Brand fixed effects, the Amazon Choice coefficient is not statistically significant and even negative. The validity of this model is, however, challenged by a low p-value in the Hansen and Sargan test for overidentifying restrictions, showing that by not controlling for brand fixed effects there might be biases in the GMM estimation. Therefore the most robust model is considered model (2).

The coefficient σ represents how much consumers value the division in nests compared to products characteristics. σ is bounded between 0 and 1 and it is zero when nests are not important, while it is 1 when characteristics are not important. The values of σ in models (1) and (2) are reasonable and show a good balance in the analysis.

7.2.2 Three Levels

The three-levels Nested Logit approach is used to express the behavioral type consumer as the one that only buys Amazon's Choice products, while the rational consumers are the ones that buy products different from Amazon's Choice. The coefficients are reported in models (3) and (4) of table 7.2. They are consistent with the ones seen for the two levels Nested Logit and are all statistically significant at the 5% or 1% level. The most relevant result here is the coefficient of $\sigma_{h,g}$, which in both models shows a high value, larger than σ_g. This shows that the division between Amazon's Choice and non Amazon's Choice can be even more important than the division among the different nests for consumers. The importance of this result comes from the fact that the products in the different nests are structurally different, therefore if $\sigma_{h,g} > \sigma_g$, this means that consumers consider Amazon's Choice products as fundamentally different from the other products.

These preliminary analyses show that the importance of Amazon's Choice goes beyond the one of a simple characteristic but is relevant enough to systematically shift consumers buying patterns

7.3 Random Coefficient Logit

The Random Coefficient Logit allows to account in a more precise way for consumer heterogeneity. Here consumers choose from the full set of products, like in the multinomial logit. Unlike the previous models, however, a set of interactions of products characteristics with demographics or standard normal variables is constructed in order to simulate a virtual sample of individuals. In the following model, two forms of heterogeneity are presented: in the first one consumers are simulated from a simple random variable v iid and normally distributed. In the second model, agents are simulated starting from a full set of demographic variables, as well as the variable v which are both interacted with products characteristics. The results of the first model are expressed in table 5 and two specifications of the model are presented: model (1) without brand fixed effects and model (2) with brand fixed effects.

In both models the Amazon's Choice coefficient is large and statistically significant, but much smaller than our previous Nested Logit estimates. The effect of the Amazon's Choice coefficient is two to three times the effect of offering the product in a bundle here. Price is negative and never statistically significant, while Rating*Reviews is positive and not statistically significant as well.

7.4 Full Model Estimates

The more interesting results come from the full model estimation, which includes demographic variables to simulate individuals. The set of demographics include $Log(Income)$, $Log(Income)^2$ Age and $Mobile$. The interactions estimated are present for all the variables,

Table 5: Random Coefficient Logit Estimates

	(1)	(2)
θ_1		
Amazon's Choice	0.182***	0.2978**
	(0.0188)	(0.118)
Price	-0.00454	-0.00800
	(0.0258)	(0.0120)
Rating*Reviews	2.483e-06	8.29e-07
	(4.61e-06)	(2.25e-06)
Bundle	0.116***	0.0906
	(0.0434)	(0.0617)
Constant	-5.634***	-5.622***
	(0.147)	(0.124)
θ_2		
v*Amazon's Choice	-8.37e-4	-7.60e-4
	(147.121)	(180.413)
v*Price	-0.00859	-0.0103
	(47.684)	(21.385)
v*Rating*Reviews	0.00296	0.00321
	(0.0133)	(0.00964)
v*Bundle	-5.11e-4	-3.21e-4
	(70.872)	(104.501)
Observations	1,059	1,059
Brand FE		Yes

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

excluding the constant, which is left out of the random component of demand. This is due to the fact that the demographic variables and the random component are not expected to have a direct impact on consumers' utility. These variables, in fact, can only affect the utility of buying a product when interacted with products characteristics.

As in the previous Sections, two models are presented, one with brand fixed effects and one without. In the estimation without brand fixed effects reported in Table 6, it is easy to see that the mean coefficients β are similar to the ones estimated in the simpler Random Coefficient model without demographic variables. however, the more precise consumer specification through demographics results in much lower standard errors for the mean effects β. The interactions with demographics are less interesting and generally not statistically significant. It is difficult to propose an explaination for the occasionally significant coefficients as there seems to be no structural meaning to these isolated cases of significant coefficients. The lack of significance in the demographics coefficients may also be due to the little number of markets considered in this work, as it is not possible to consider more than the five European Amazon markets. In the staple studies that use the random coefficient logit approach, such as Nevo (2001), the number of markets is much larger (in that case 94 markets were considered), allowing for more differentiation among the distributions of demographics in different markets.

In Table 7 the results of the full model with demographics and brand Fixed Effects are reported. Here as well most of the interactions are not statistically significant, and the mean coefficients are estimated with higher precision than in the model without demographics (5). All the previous conclusions on mean effects hold here as well, with Amazon's Choice having a strong positive and statistically significant effect on Utility of the agents and likely driving a large portion of the demand. It is important to note a strong and statistically significant interaction of Rating*Reviews with $Log(Income)^2$ and $Mobile$ which shows that Rating*Reviews is likely more relevant for high income groups and for people using mobile phone to shop online. Bundle also shows a positive and statistically significant interaction with $Log(Income)$ even thogugh the size of the interaction is small when compared with the mean effect of Bundle on consumers utility. This suggests that higher income agents are slightly more likely to buy a bundle of products to store for later use than to buy a single product for a one-time use.

Table 6: Full Model, No Brand FE

	Mean Effects	Random	Demographics			
	β	v	$Log(Income)$	$log(Income^2)$	Age	$Mobile$
Amazon's Choice	0.183***	-0.00172	1.29e-4	0.00129	0.0208**	0.0156
	(3.15e-6)	(0.00242)	(0.0129)	(0.0129)	(0.00940)	(0.0597)
Price	-0.00454**	-0.00151	8.91e-4	-0.0238	0.00489	-0.00133
	(0.00198)	(0.915)	(0.836)	(0.836)	(0.471)	(1.41)
Rating*Reviews	2.52e-6	0.00293	5.68e-4	0.0408***	8.58e-4	0.0309*
	(3.623e-5)	(0.101)	(0.0113)	(0.0113)	(0.0382)	(0.0176)
Bundle	0.116***	3.53e-5	-0.00104	3.26e-4	7.12e-4	-0.0361
	(6.08e-5)	(0.0417)	(0.0295)	(0.0295)	(0.00771)	(0.0400)
Constant	-5.635***					
	(1.93e-6)					

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Full Model, With Brand FE

	Mean Effects	Random	Demographics			
	β	v	$Log(Income)$	$log(Income^2)$	Age	$Mobile$
Amazon's Choice	0.299***	-0.00165	-1.43e-4	0.00133	0.0205	0.0164
	(0.0563)	(73.325)	(73.969)	(73.969)	(39.222)	(80.302)
Price	-0.00800	-0.00161	-7.35e-4	-0.0256	0.00877	-0.00139
	(0.0275)	(56.469)	(3.884)	(3.884)	(20.883)	(13.945)
Rating*Reviews	8.82e-7	0.00316	5.24e-4	0.0429***	8.36e-4	0.0327*
	(2.30e-5)	(0.0255)	(0.00293)	(0.00293)	(0.0314)	(0.0181)
Bundle	0.0906*	-6.91e-5	1.52e-4***	2.33e-4	6.43e-4	-0.0339
	(0.0563)	(37.612)	(65.361)	(65.360)	(25.547)	(58.907)
Constant	-5.622***					
	(0.105)					

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

7.5 Results interpretation

The presence of such large coefficients across different models and estimation techniques suggests that the Amazon's Choice badge has a substantial effect on consumers' demand. Understanding the exact process through which the badge acts is out of the scope of this paper, but it is recognized here that there are multiple interpretations of the Amazon's Choice coefficient.

The rational model presented follows the work of Hui et al. (2016) and Tadelis (2016) in interpreting the Amazon's Choice badge as a signal of seller reputation and product quality. In this case, consumers consider all possible options, but value more the Amazon's Choice one as it is directly suggested by the platform. This perception is not necessarily correct as it is unclear to customers how the Amazon's Choice badge is assigned.

The behavioral approach, instead, focuses on the Amazon's Choice as an attention grabber or a salient characteristic. This interpretation follows the work of Bordalo et al. (2013) in which a model of firms competition with salient characteristics is presented. The Amazon's Choice badge can in fact easily capture consumers' attention in a situation of choice overload, in which simpler options are often preferred by customers, as shown in the experiment proposed in Iyengar & Kamenica (2010).

An alternative interpretation of the coefficient could be that the time required for the transaction is not null, as it was assumed in the theoretical model of bounded rationality proposed (Section 4.1), and therefore search costs are not equal to zero. Consumers may be choosing under time constraints and the loss caused by choosing a non optimal predetermined option may be less than the value of time needed to compare all possible alternatives. Some customers may in fact prefer a predetermined option in a market such as the one analyzed, where products are similar in price and characteristics. The small benefits of a complete search for the best product may be out weighted by the time saved in purchasing the suggested option. This interpretation is consistent with the latest research on consumer's choice under time pressure. Reutskaja et al. (2011) shows that customers, while still optimizing under choice overload and time pressure, often display behavioral biases like systematically putting more attention to products displayed in the central area of their vision. Also the more recent Coey et al. (2016) shows that under time constraints consumers may settle for non-optimal choices and more expensive products.

The relevance of the interaction of Amazon's Choice and customers' age in the full BLP model without fixed effects shows a potential heterogeneous effect of the Amazon's Choice badge on consumer's demand. This heterogeneity is insignificant when brand fixed effects are included. However, future research with a larger number of markets and more precise customer data could show more relevant interactions of the Amazon's Choice badge with individual consumer's characteristics. This small heterogeneity in the effect of the Amazon's Choice badge could be due to the heterogeneous effects of the badge on consumers attention (which could vary with age); alternatively it could be due to heterogeneity in consumers' value of time. The differential effects of the Amazon's Choice badge on time should however be reflected in the interaction with income: agents with larger income should value more the Amazon's Choice option as the value of time saved during the choice should be larger. However, the interaction with income doesn't show any relevant effects on consumer demand, showing that this type of heterogeneous effect is unlikely to be present in the proposed setting.

8 Conclusions

This study aims at showing the impact of a specific Amazon Marketplace suggestion on consumer demand. To do so a simple theoretical model of discrete choice is constructed and further developed in two variations. At first all consumers are assumed to be perfectly rational and the Amazon's Choice badge is analyzed as a standard characteristic of the product. Then, a behavioral adaptation of the model is also proposed, with a fraction of the consumers only considering the Amazon's Choice product and ignoring all other available choices. The scope of this work is to find robust evidence of the influence of Amazon's Choice badge on consumers demand in both settings.

A set of structural empirical models is presented to estimate the demand function of consumers and strong evidence of the influence of Amazon's Choice badge on consumers' utility and demand is found. Starting from the full rationality approach, the structural empirical techniques used are taken from the literature of applied industrial organization and followed mainly the work of S. T. Berry (1994), S. Berry, Levinsohn, & Pakes (1995) and Nevo (2001). The techniques used are Multinomial Logit, Nested Logit and Random Coefficient Logit. These three have a gradually more precise characterization of simulated individual agents across the five European Amazon Marketplaces (Italy, United Kingdom, Spain, France, Germany). The estimates of Amazon's Choice impact for the rational coefficient range from a more conservative 0.299 in the Random Coefficient model with brand Fixed Effects, to a more extreme 1.174 in the Nested Logit with brand Fixed Effects.

These results show not only a strong and statistically significant impact of the Amazon's Choice badge on consumer utility, but also a very large impact when compared to the other products' characteristics. Among the characteristics included in the analysis (Price, Rating*Reviews and Bundle), none of them is able to rival the impact of Amazon's Choice, with Bundle having the largest effect ranging from 0.0906 to 0.22. These results suggest that the effect of Amazon's Choice on these particular set of products (highly substitutable with generally low variance in prices) is more than offering the same product in a bundle, which likely drastically increases the costs of production for the product.

Furthermore, a second set of coefficients is estimated through the use of a three-levels Nested Logit. This empiriacl setting allows to split the simulated set of consumers into two types, one that follows blindly the Amazon's Choice suggestion and another that completely ignores the Amazon's Choice badge and compares all other products. After the division of products into different nests (common to any Nested Logit model), consumers are also divided in agents which are likely to buy Amazon's Choice and consumers who only consider the other products. The approach presented allows to estimate the significance of this further division in the form of the coefficient $\sigma_{h,g}$. This coefficient can theoretically range from 0 to 1 and in our estimation is 0.498 without brand fixed effects and 0.590 with brand fixed effects. The theoretical explanation of the coefficient $\sigma_{h,g}$ suggests that it is closer to 0 when the division is meaningless, while it is closer to 1 when the division is the only meaningful determinant of demand, and products characteristics don't have any impact on products demand. The estimates presented show that the division between Amazon Choice products and non Amazon Choice products is likely more important then all other characteristics which define each product. Showing not only that the Amazon's Choice badge has an impact stronger than all other characteristics taken individually, but the impact is likely stronger than the combined effect that all these different characteristics can have on demand.

All these estimates together show how strong the effect of a suggestion badge can be in online platforms. The purpose of the work is in fact not to draw conclusions on the channels through which the suggested choices act in consumers' decisions. Instead this works proves that the Amazon's Choice badge has a strong and statistically significant effect on consumers' decisions, through the use of different theoretical models and empirical techniques.

This work contributes to the existing literature of Empirical Industrial Organization in showing that even in an environment with almost null search costs (such as a common online platform with similar products), the role of suggestions and attention grabbers seems to be fundamental in consumers' choices. The results proposed here are likely driven by the peculiarity of the market presented, in which single products characteristics may not be as relevant as in markets with highly differentiated products. Nonetheless the results shown here show that the platform itself is capable of driving demand to the desired products with very little effort (simply suggesting a product over others in a noninvasive and subtle way).

The implications of these results are even more relevant due to the fact that Amazon does not reveal how the Amazon's Choice badge is assigned, leaving sellers to speculate how to reach it. A non-neutral environment for the assignment of the Amazon's Choice badge, which could favour certain sellers over others, is likely to have a dramatic impact on the relevant market for such product, driving demand away or towards certain sellers (among which also Amazon competes either producing products under the “Amazon Basics” brand or buying products from producers and selling them directly).

This work suggests for more transparency on the assignment of the Amazon’s Choice badge and other instruments similar to this one. Under the current European legislation (Council of the European Union, 2019), online platforms are required to disclose the main determinants of ranking algorithms, among which, by article 2.8, Amazon’s Choice should be considered. It remains to be seen whether a caveat on consumer protection and personalization of results (article 5.6) is used to both protect both the consumers data and the intellectual property of the algorithm determining the Amazon’s Choice products. The more recent “New Competition Tool” (European Commission, 2020b) and the “Digital Service Act Package” (European Commission, 2020a) initiative, which both aim at being adopted by the end of 2020, propose specific tools for the control of large internet platforms that act as gatekeepers and intermediaries. Among the objectives expressed in section B of the “Digital Service Act Package”, at point 2 is written: “[...] further horizontal rules could be envisaged with a purpose to enable collection of information from large online platforms acting as gatekeepers by a dedicated regulatory body at the EU level to gain, for example, further insights into their business practices and their impact on these platforms’ users and consumers”. It remains to be seen whether the proposed tools will have the power to assess the neutrality of the assignment process of the Amazon’s Choice badge and similar platform suggestions.

References

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, *58*(2), 277–297.

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, *37*(3), 668–691.

Ben-Akiva, M., & Bierlaire, M. (1999). Discrete choice methods and their applications to short term travel decisions. In *Handbook of transportation science* (pp. 5–33). Springer.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, *121*(5), 803-843.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2015). Competition for Attention. *The Review of Economic Studies*, *83*(2), 481-513.

Brenkers, R., & Verboven, F. (2006). Liberalizing a distribution system: the european car market. *Journal of the European Economic Association*, *4*(1), 216–251.

Bronnenberg, B. J., Dubé, J.-P., Gentzkow, M., & Shapiro, J. M. (2015). Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium. *The Quarterly Journal of Economics*, *130*(4), 1669-1726.

Brown, Z. Y., & Jeon, J. (2020). *Endogenous information and simplifying insurance choice.* Mimeo, University of Michigan.

Caillaud, B., & Jullien, B. (2003). Chicken & egg: Competition among intermediation service providers. *RAND journal of Economics*, 309–328.

Cardell, N. S. (1997). Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory*, 185–213.

Coey, D., Larsen, B., & Platt, B. (2016). Discounts and deadlines in consumer search. *National Bureau of Economic Research*.

Council of the European Union. (2019). *Council regulation (EU) no 1150/2019.* Retrieved from `https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R1150&from=EN` (Last access 10 November, 2020)

Decarolis, F. (2015). Medicare part d: Are insurers gaming the low income subsidy design? *American Economic Review*, *105*(4), 1547–80.

Eliaz, K., & Spiegler, R. (2011). On the strategic use of attention grabbers. *Theoretical Economics*, *6*, 127 - 155. doi: 10.3982/TE758

Ellison, G. (2006). Bounded rationality in industrial organization. In R. Blundell, W. K. Newey, & T. Persson (Eds.), *Advances in economics and econometrics: Theory and applications, ninth world congress* (Vol. 2, p. 142–174). Cambridge University Press.

Ellison, G., & Ellison, S. F. (2005). Lessons about markets from the internet. *Journal of Economic Perspectives*, *19*(2), 139-158.

European Commission. (2019). *Antitrust: Commission opens investigation into possible anti-competitive conduct of amazon.* Retrieved from `https://ec.europa.eu/commission/presscorner/detail/en/IP_19_4291` (Last access 10 November, 2020.)

European Commission. (2020a). *Digital service act package ex ante regulatory instrument of very large online platforms acting as gatekeepers.* Retrieved from `https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/`

12418-Digital-Services-Act-package-ex-ante-regulatory-instrument-of-very-large-online-platforms-acting-as-gatekeepers (Last access 10 November, 2020. Ref: Ares(2020)2877647 - 04/06/2020)

European Commission. (2020b). *New competition tool - inception impact assessment.* Retrieved from https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12416-New-competition-tool (Last access 10 November, 2020. Ref: Ares(2020)2877634 - 04/06/2020)

European social survey round 9 (Data file edition 2.0 ed.). (2018). doi:10.21338/NSD-ESS9-2018. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

Eurostat. (2020). *Distribution of income by quantiles - eu-silc and echp surveys (ilc_di01).* Retrieved from https://ec.europa.eu/eurostat/data/database (Last access 10 November, 2020)

Gandhi, A., Lu, Z., & Shi, X. (2013). *Estimating demand for differentiated products with error in market shares.* Cemmap working paper.

Goldberg, P. K., & Verboven, F. (2001). The evolution of price dispersion in the european car market. *The Review of Economic Studies*, *68*(4), 811–848.

Grubb, M. D. (2015). Failing to choose the best price: Theory, evidence, and policy. *Review of Industrial Organization*, *47*(3), 303-340.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.

Hortaçsu, A., Madanizadeh, S. A., & Puller, S. L. (2017). Power to choose? an analysis of consumer inertia in the residential electricity market. *American Economic Journal: Economic Policy*, *9*(4), 192–226.

Hui, X., Saeedi, M., Shen, Z., & Sundaresan, N. (2016). Reputation and regulations: evidence from ebay. *Management Science*, *62*(12), 3604–3616.

Iyengar, S. S., & Kamenica, E. (2010). Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics*, *94*(7-8), 530–539.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, *69*(2), 307–342.

Pilello, A., & De Cristofaro, B. (2020). Report annuale sull'e-commerce italiano: come sta evolvendo il commercio digitale nel nostro paese. Retrieved from `https://www.idealo.it/magazine/2020/04/02/e-commerce-italia-report-2020-idealo/` (Last access 10 November, 2020)

Prat, A., & Valletti, T. M. (2019). *Attention oligopoly.* SSRN. Retrieved from `https://ssrn.com/abstract=3197930` (Last access 10 November, 2020)

Radner, R. (1975). *Optimization techniques ifip technical conference.*

Rasmusen, E., et al. (2007). *The blp method of demand curve estimation in industrial organization.* Department of Business Economics and Public Policy, Kelley School of Business, Indiana University. Retrieved from `https://www.rasmusen.org/published/blp-rasmusen.pdf` (Last access 10 November, 2020)

Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review*, *101*(2), 900–926.

Rochet, J.-C., & Tirole, J. (2006). Two-sided markets: a progress report. *The RAND journal of economics*, *37*(3), 645–667.

Saeedi, M. (2014). Reputation and adverse selection, theory and evidence from ebay. *Theory and Evidence from eBay*.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica: Journal of the Econometric Society*, 393–415.

Shifflett, S., Berzon, A., & Mattioli, D. (2019). Amazon's choice isn't the endorsement it appears. *The Wall Street Journal*. Retrieved from `https://www.wsj.com/articles/amazons-choice-isnt-the-endorsement-it-appears-11577035151` (Last access 10 November, 2020)

Smallwood, D. E., & Conlisk, J. (1979). Product quality in markets where consumers are imperfectly informed. *The Quarterly Journal of Economics*, *93*(1), 1–23.

Spiegler, R. (2011). Bounded rationality and industrial organization. Oxford University Press.

Stahl, D. O. (1989). Oligopolistic pricing with sequential consumer search. *The American Economic Review*, 700–712.

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, *8*, 321–340.

Train, K. E. (2009). *Discrete choice methods with simulation.* Cambridge university press.

Tversky, A. (1974). Assessing uncertainty. *Journal of the Royal Statistical Society: Series B (Methodological)*, *36*(2), 148-159.

Varian, H. R. (1980). A model of sales. *The American Economic Review*, *70*(4), 651–659.

Zervas, G., Proserpio, D., & Byers, J. W. (2017). The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry. *Journal of marketing research*, *54*(5), 687–705.

Appendix

Three-Levels Nested Logit

The purpose of this appendix is to provide the derivation of the three-levels Nested Logit closed form. It is constructed starting from the demonstration of the two-levels Nested Logit provided in S. T. Berry (1994) and from the insights of Brenkers & Verboven (2006).

As usual in all the nested logit models, the first step is to write a simple utility function for the customer i buying product j.

$$\mu_{i,j} = \mathbf{x}_j'\boldsymbol{\beta} - \alpha p_j + \varepsilon_j + \epsilon_{i,j} \tag{44}$$

Where $\epsilon_{i,j}$ accounts for consumer heterogeneity and is assumed IID with Extreme Value type 1 distribution. Assuming each customer will only buy one product, the quantity of product j sold is a linear function of the probability that a customer will buy product j:

$$q_j = M * p_j(\mathbf{x}, \boldsymbol{\varepsilon}, \mathbf{p}, \theta) \tag{45}$$

Where M indicates market size. Therefore it is easy to write the predicted market share as:

$$d_j = q_j/M = p_j(\mathbf{x}, \boldsymbol{\varepsilon}, \mathbf{p}, \theta) \tag{46}$$

It is shown by Brenkers & Verboven (2006) that in the three level nested logit model, this can be rewritten as:

$$d_j = d_{j|h,g} * d_{h|g} * d_g \tag{47}$$

Where d_j is the predicted market share of product j, $d_{j|h,g}$ is the predicted market share of j in the subnest h and group g, $d_{h|g}$ is the predicted market share of subnest h inside group

g, d_g is the predicted market share of group g among all groups $g \in G$. This is equal to:

$$d_j = \frac{exp\{\delta_j/(1-\sigma_{hg})\}}{exp\{\ell_{ihg}/(1-\sigma_{hg})\}} * \frac{exp\{\ell_{ihg}/(1-\sigma_g)\}}{exp\{\ell_{ig}/(1-\sigma_g)\}} * \frac{exp\{\ell_{ig}\}}{exp\{\ell_i\}} \tag{48}$$

where:

$$\delta_j = \mathbf{x}_j'\boldsymbol{\beta} - \alpha p_j \tag{49}$$

and:

$$\begin{aligned} \ell_{ihg} &= (1-\sigma_{hg}) \ln \sum_{j \in h} exp\{\delta_j/(1-\sigma_{hg})\} \\ \ell_{ig} &= (1-\sigma_g) \ln \sum_{h \in g} exp\{\ell_{ihg}/(1-\sigma_g)\} \\ \ell_i &= \ln \sum_{g \in G} exp\{\ell_{ig}\} \end{aligned} \tag{50}$$

Equation (48) is rewritten for demonstration purposes, defining:

$$\begin{aligned} D_{hg} &= \sum_{j \in h} exp\{\delta_j/(1-\sigma_{hg})\} \\ D_g &= \sum_{h \in g} D_{hg}^{\frac{1-\sigma_{hg}}{1-\sigma_g}} \end{aligned} \tag{51}$$

therefore

$$d_j = \frac{exp\{\delta_j/(1-\sigma_{hg})\}}{D_{hg}} * \frac{D_{hg}^{\frac{1-\sigma_{hg}}{1-\sigma_g}}}{D_g} * \frac{D_g^{1-\sigma_g}}{\sum\limits_{g \in G} D_g^{1-\sigma_g}} \tag{52}$$

The outside good is defined as a separate group with mean utility $\delta_0 = 0$, therefore:

$$d_0 = \frac{1}{\sum\limits_{g \in G} D_g^{1-\sigma_g}} \tag{53}$$

By taking the ratio of (52) and (53):

$$\begin{aligned} d_j/d_0 &= \frac{exp\{\delta_j/(1-\sigma_{hg})\}}{D_{hg}} * \frac{D_{hg}^{\frac{1-\sigma_{hg}}{1-\sigma_g}}}{D_g} * D_g^{1-\sigma_g} \\ &= \frac{exp\{\delta_j/(1-\sigma_{hg})\}}{D_{hg}^{\frac{\sigma_{hg}-\sigma_g}{1-\sigma_g}} * D_g^{\sigma_g}} \end{aligned} \tag{54}$$

And by taking the previous equation in logs:

$$\ln(d_j) - \ln(d_0) = \frac{\delta_j}{1-\sigma_{hg}} - \frac{\sigma_{hg}-\sigma_g}{1-\sigma_g} \ln(D_{hg}) - \sigma_g \ln(D_g) \tag{55}$$

From (47) and (52) it is easy to see that:

$$\begin{aligned} d_{j|h} &= \frac{D_{hg}^{\frac{1-\sigma_{hg}}{1-\sigma_g}}}{D_g} \\ d_{h|g} &= \frac{D_g^{1-\sigma_g}}{\sum\limits_{g \in G} D_g^{1-\sigma_g}} \end{aligned} \tag{56}$$

It is trivial to show then that:

$$\ln(d_j) - ln(d_0) - \sigma_{hg} \ln(d_{j|h}) - \sigma_g \ln(d_{h|g}) = \delta_j \tag{57}$$

By replacing the predicted market shares (d) with the observed market shares(S), a closed form solution is obtained, in the form of the following econometric equation:

$$\ln(S_j) - \ln(S_0) = \mathbf{x}_j'\boldsymbol{\beta} - \alpha p_j + \sigma_{hg}\ln(S_{j|h}) - \sigma_g \ln(S_{h|g}) + \varepsilon_j \tag{58}$$

Where σ_g and σ_{hg} are restricted between 0 and 1 as shown in Cardell (1997). The closer they are to 0, the more the variation within nest is important for consumers (if they are both equal to 0, we are back at the simple logit model), the closer they are to 1 the more the variation between nests is important for consumers (Train, 2009).