# Entry Into Two-Sided Markets Shaped By Platform-Guided Search

Kwok Hao Lee, Leon Musolff

September 30, 2021

Click Here For Most Recent Version

**Abstract**

We evaluate the problem of firms that operate platforms matching buyers and sellers, while also selling goods on these same platforms. By being able to guide consumer search through algorithmic recommendations, these firms can influence market outcomes, a finding that has worried regulators. To analyze this phenomenon, we combine rich novel data about sales and recommendations on Amazon Marketplace with a structural model of intermediation power. In contrast to prior literature, we explicitly model seller entry. This feature enables us to assess the most plausible theory of harm from self-preferencing, i.e. that it is a barrier to entry. We find that recommendations are highly price elastic but favor Amazon. A substantial fraction of customers only consider recommended offers, and recommendations hence noticeably raise the price elasticity of demand. By preferring Amazon's offer, the recommendation algorithm raises consumer welfare by approximately $4.5 billion (since consumers also prefer these offers). However, consumers are made worse off if self-preferencing makes the company raise prices by more than 7.8%. By contrast, we find no evidence of consumer harm from self-preferencing through the entry channel. Nevertheless, entry matters. The algorithm raises consumer welfare in the short and medium run by increasing the purchase rate and intensifying price competition. However, these gains are mostly offset by reduced entry in the long run.

[1] Princeton University, Department of Economics. Lee: `khl@princeton.edu`; Musolff (corresponding author): `lmusolff@princeton.edu`. The authors are indebted to Jakub Kastl, Kate Ho, Nicholas Buchholz, Adam Kapor and Bo Honoré for invaluable guidance. We would also like to thank Bruno Baránek, Michael Dinerstein, Jean-François Houde, Ian Kaneshiro, Uri Miron and members of the European Commission for their insightful feedback. We are also grateful for the support from the Department of Economics at Princeton University and our data provider.

# 1 Introduction

## 1.1 Motivation

This paper examines the interplay between Amazon's need to attract merchants and its ability to guide consumer search. The platform faces a key tradeoff between fostering competition and providing incentives for these merchants to enter. This tradeoff is complicated by the company's participation on its platform as a seller. As profits from sales are an attractive substitute to intermediation fees, economists worry that platforms in Amazon's position may be tempted to self-preference, i.e., guide consumers preferentially towards their own offers (Piontek 2019). This concern has attracted attention from antitrust agencies (Chunduru 2020) and regulators (U.S. Congress 2021; U.S. Senate Judiciary Committee 2020). In addition to static harm from not showing consumers their most preferred offers, proposed theories of harm from self-preferencing focus on a possible stifling of merchant entry and associated reduction in options available to consumers (Klingler et al. 2020).

Even when firms refrain from participating as merchants on their platforms, their ability to steer consumer choice raises economic questions about the extent of their influence on market outcomes. However, rigorous empirical analysis has lagged popular attention because entry models are computationally complex and the relevant data are scarce.

We aim to close this gap in two ways. Firstly, we build a tractable structural model of intermediation power, i.e., a platform's ability to influence market outcomes by steering consumers. We model demand, the platform's choice of recommendations (which guide consumer search), pricing, and (our key contribution) entry. Explicitly accounting for entry allows our model to speak directly to the dynamic theory of harm from self-preferencing and the tradeoff between attracting merchants and fostering competition. Secondly, we combine our model with extensive novel proprietary data on sales and search guidance decisions on Amazon, a setting of prime regulatory concern.

Our paper builds on the observation that platforms commonly set defaults for consumers through their search, ranking, and recommendation algorithms. As economists have increasingly become aware[2], these defaults can influence consumer choice (Thaler and Sunstein 2008). We thus model consumers as having one of two kinds of consideration sets (Goeree 2008): either they consider all options available to them, or they only consider the recommended offers and the outside option. Through its influence on consideration sets, the platform's recommendation algorithm affects seller pricing, as well as *how many* and *which* firms enter.

We estimate our model on high-frequency data from Amazon, a critical real-world example of a firm that participates as a merchant on its own e-commerce platform. We find that the firm indeed has intermediation power: 26% of consumers only consider offers recommended by the platform, thereby allowing the algorithm to raise the average price elasticity of demand from 3 to 4. Furthermore, it uses this power to intensify price competition (recommendations load heavily

[2]As have firms: e.g., in 2018, Google paid Apple $12B to remain Safari's default search engine (Segarra 2018).

on price with an elasticity of 19) and preferentially guide consumers towards its own offers (which have an advantage equivalent to a 9% price discount).

After recovering recommendation, demand, and cost parameters, we perform two counterfactual analyses to evaluate different designs of the recommendation algorithm. In these analyses, we proceed in three steps. Firstly, we compare the short-term impacts of a change in the algorithm, keeping prices and entry fixed. This step reflects the results platform companies would find when running brief experiments ("A/B tests") on their algorithms, as is standard industry practice. Secondly, we allow merchants to adjust their pricing decisions in the medium-run scenario. Finally, our innovation on the literature lies in our modeling of entry decisions, which we allow to vary in the long-run scenario.

We start by assessing the equilibrium effects of self-preferencing, i.e., Amazon recommending its own offers at an elevated frequency. In the short run, we find that self-preferencing raises consumer welfare as consumers themselves prefer Amazon's offers, perhaps due to them being less likely to be counterfeit. However, the short-run gains are offset by 0.5% higher prices and a 0.01 reduction in entrants per market in the medium and long run. Nevertheless, the considerable static welfare gains outweigh the dynamic impact. Overall, the platform's recommendation advantage increases consumer welfare by $4.5 billion. However, we show in a supplemental analysis that if Amazon were to raise prices by more than 7.8% due to self-preferencing, the practice could nevertheless harm consumers. Nevertheless, our model does not validate the theory that self-preferencing harms consumer welfare by acting as a barrier to entry.

The combined effect of search guidance on welfare is also positive. Consumers enjoy a static welfare gain of $50.9 billion from Amazon's recommendation algorithm. The intuition behind this result is that offers on online marketplaces can vary dramatically in their attractiveness, and guiding consumers towards attractive offers can thus be very helpful. However, these gains rely on the recommendation algorithm's emphasis on price, which causes a 14.2% price decrease in the medium run. This price decrease is a redistribution of surplus from the producer to the consumer side of the economy (which enjoys an overall welfare gain of $60.7 billion in the medium run). As such, it is not without consequences. Indeed, in the long run, the presence of recommendations causes a reduction in entry of 2.21 entrants/market. This reduction dramatically lowers the overall consumer welfare gains from search guidance technology to $0.9 billion. Thus, while benefiting naive consumers (who follow the recommendation and therefore gain little from the availability of multiple offers), the recommendation algorithm can potentially harm sophisticated consumers (who would have exploited their wealth of choices).

## 1.2 Literature Review

This paper speaks to four strands of literature in Industrial Organization and the Economics of Digitization.

Firstly, we contribute to a large literature on price competition and search frictions in online markets. Despite theoretically small search costs, online markets still see substantial price dispersion (Bailey 1998; Smith and Brynjolfsson 2001; Baye, Morgan, and Scholten 2004; Einav et al. 2015), possibly because consumers do not search efficiently (Malmendier and Lee 2011; Stigler 1961) and consider only a subset of the available alternatives (Goeree 2008). In the presence of these frictions, platforms face a trade-off between incentivizing sellers to compete on price and guiding consumers to their preferred products (Dinerstein et al. 2018). With an efficient search technology (e.g., a product search engine), offer-level elasticities can be very high; in response, retailers may obfuscate product characteristics to hinder comparison and raise profits (Ellison and Ellison 2009). Obfuscation has intrigued theorists, spawning a literature on optimal recommender systems (Che and Hörner 2017) and how platforms should curate their sellers (Casner 2020). While focusing on similar platform-design problems as us, these papers do not feature a company competing on its own platform and do not model the entry channel.

Secondly, while the interplay between pricing and entry is novel to the platform-guided search literature, the importance of additional draws from the entrant distribution is commonly appreciated in the literature on the value of variety. Online markets benefit from having essentially "infinite shelf space", raising product variety in many markets, including those for books (Brynjolfsson, Hu, and Smith 2003). With improvements in recommender systems, consumers enjoy the "long tail" of products offered by niche sellers (Brynjolfsson, Hu, and Smith 2006; Donnelly, Kanodia, and Morozov 2021). However, taste heterogeneity is not needed for additional sellers to benefit consumers as marginal entrants may be of high quality when such quality is ex-ante hard to predict (Aguiar and Waldfogel 2018). For instance, after GDPR[3], consumer surplus was decreased through precisely this channel as fewer new apps entered the Google Play Store (Janssen et al. 2021).

Thirdly, we build on the literature on estimating entry costs (Bresnahan and Reiss 1991; see Berry and Reiss 2007 for a survey). Entry costs have been estimated in the markets for air travel (Berry 1992) and motels (Mazzeo 2002). Multiple equilibria are a common threat to point-identification in these papers. The usual solution is partial identification (Tamer 2003; Ciliberto and Tamer 2004; Chernozhukov, Hong, and Tamer 2007). We instead assume that sellers play a symmetric equilibrium, an approximation to a possibly more complicated dynamic model. We believe this assumption is innocuous as our setting is characterized by high turnover and the presence of small, niche sellers on long-tail products.

Finally, our paper relates to the literature on selective entry into auctions (see Hendricks and Porter 2007 for a survey). In these papers, bidders decide whether to pay a fixed cost to enter the auction; if they enter, more private information may be revealed to them, and play proceeds amongst the entrants as in a typical auction game. Historically, in the entry stage, it has been assumed that bidders either know nothing about their value for the good(s) to be auctioned off (Levin and Smith 1994) or know it exactly (Samuelson 1985). Our paper is closest to those of Roberts and Sweeting (2010, 2013), in which the authors make an intermediate informational assumption: each potential entrant receives a signal about her value before she decides whether

[3] General Data Protection Regulation

to enter. Our equilibria, like theirs, can be characterized as cutoff values in the agents' respective signals: an agent enters if and only if her signal says that entry will yield positive expected profits.

## 2 Setting

Our empirical setting is Amazon Marketplace, a platform permitting third parties to list their offers among Amazon's. In 2020, about 1.9 million merchants exploited this opportunity (Bezos 2021). Their listings were responsible for 60% of retail sales, yielding total estimated merchant profits of $25 billion (Bezos 2021). However, with 28% of purchases on Amazon completed in three minutes or less, customers seem to use little time to explore their (extensive) options (Bezos 2021). This speed suggests that search, ranking, and recommendation algorithms play an outsized role in shaping consumers' choices.

There are multiple such algorithms on any e-commerce platform. However, when products differ on unobserved dimensions, investigating recommendation algorithm behavior is challenging. Therefore, while prior literature restricts attention to a particular product (e.g. Dinerstein et al. 2018), we exploit a unique feature of our chosen e-commerce platform to sidestep the issue. Although many platforms (e.g., eBay) do not distinguish between multiple offers on the same product and offers for different products, Amazon does make this distinction. It requires sellers to list their offers on the correct product page — even if this means that a specific product may have multiple offers.

When there are multiple offers, the platform automatically designates *at most one* of the offers as "recommended." Amazon's recommendation is vital to sellers because the recommended offer is placed in the coveted "Buybox", as depicted in Figure 1. Merchants and other market participants alike know that the "vast majority of sales are done through" the Buybox (Caudet and Tsoni 2019). Indeed, "industry experts estimate that about 80% of Amazon sales go through the Buy Box" (U.S. Senate Judiciary Committee 2020). Naturally, therefore, sellers seek to understand how the recommendation algorithm behaves. Its inner workings are a subject of rife speculation on online message boards. Indeed, one of the leading performance metrics made available to merchants is their "Buy Box Percentage", reflecting the fraction of product views for which a merchant was recommended.

To the econometrician, this algorithm provides an immaculate setting in which to examine the impact of platform-guided consumer search. In particular, a market will be a specific product — e.g. "Clarks Men's Bushacre 2 Chukka [Shoes], Dark Brown, Size 8.5." Once they have decided on this product, consumers must decide between various offers. The merchant "Zappos", for instance, offers these shoes for US$57.68 and will deliver them via Amazon Prime. Alternatively, "BHFO" will charge only US$54.99 but does not offer Prime shipping. Below, we will examine the recommendation's influence on how consumers choose between these offers. Crucially, in this choice, all options share all product characteristics. This feature of the setting obviates the need to estimate complex demand models to match the observed substitution patterns.

Finally, Amazon Marketplace is also an important setting to evaluate antitrust concerns. Economists worry that "defaults can direct a consumer to the choice that is most profitable to the platform" (Piontek 2019). Indeed, sellers complain that it can be challenging to compete on products that Amazon sells as the platform will frequently assign the Buybox to itself. A Senate investigation into this practice concluded that "Amazon can give itself favorable treatment relative to competing sellers. It has done so through its control over the Buy Box" (U.S. Senate Judiciary Committee 2020). The investigation sparked the introduction into Congress of a draft bill prohibiting, "in connection with any user interfaces, including search or ranking functionality offered by the covered platform, treat[ing] the covered platform operator's own products, services, or lines of business more favorably than those of another business user" (U.S. Congress 2021).

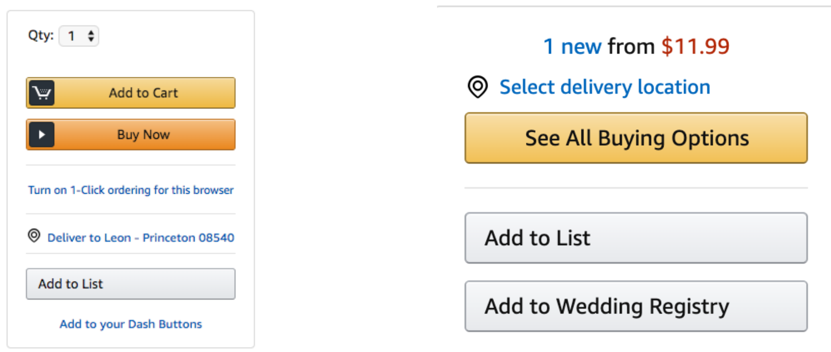
Recently, there have also been allegations that the company tweaked its search algorithms to give preferential treatment to products sold by its own retail unit (Mattioli 2019). These allegations suggest the sellers' worries about the Buybox may be justified. However, when speaking to the Senate, Amazon's general counsel Nate Sutton emphasized that "the Buy box is aimed to predict what customers want to buy" and that "[the platform applies] the same criteria whether [the merchant is] a third-party seller or Amazon" (Neidig 2019). Nevertheless, antitrust authorities across the world continue to investigate this issue (Caudet and Tsoni 2019).

### 2.1 Data

We procured extensive, high-frequency data that is novel to this literature. We see prices, recommendations, and sales on 200,000 products across about 1,000 sellers over 1.5 years. Our data are sourced from a company that offers "repricing" services; therefore, these data arise directly from Amazon's APIs. As a result, our observations are of much higher fidelity and frequency than comparable web-scraped data. To be exact, for each of the products that the repricing company monitors, Amazon automatically notifies the company when there is any change in the number or content of the offers (e.g., a price hike or the entry of a new competitor). Importantly, we see which offer Amazon recommends *just after* any such change. However, these recommendations are in flux, and the company is not sent a notification if only the recommendation status has changed.

A distinctive feature of our data is that we observe sales for exactly one merchant on each market (the merchant employing the services of the repricing company). While these data come with their own challenges (such as having to proxy for the market size), to our knowledge, no past papers have employed sales data across a large number of merchants to credibly estimate demand on Amazon Marketplace. In the absence of such data, it has become common to employ sales ranks as a proxy. These ranks, however, aggregate sales at the product-page level. Thus, they do not allow the econometrician to uncover the relationship between the recommendation status of an offer and its market share.

We discuss our data in more detail in Appendix A.




**(a)** *Occupied Buybox.* **(b)** *Canceled Buybox.*

**Figure 1:** *The Buybox.*

Notes*: The "Amazon Buybox", through which most sales on Amazon are made. Amazon chooses which seller is assigned the sale if the buttons inside the rectangle are used. In (a), a seller has been assigned while in (b), the Buybox has been deliberately left empty.*

### 2.2 Canceled Recommendations

Before we delve into the structural model, we show some evidence that suggests that the platform cares about the *equilibrium* implications of its recommendation algorithm design. In particular, we are interested in a curious phenomenon: sometimes, the platform makes no recommendations on a particular product. In such a case, the consumer is shown the Buybox in Figure 1b.

Prima facie, this behavior is counterintuitive: if the platform does not recommend any offer, it should see lower sales in the short run. This is because customers who rely only on the recommendation do not buy anything. As a result, the merchants on this product are less likely to make a sale, and the platform is less likely to earn its intermediation fees. Indeed, we confirm in Table 1 that, controlling for product fixed-effects and prices, a canceled recommendation is indeed associated with a 26.31% lower daily sales probability. However, in the medium run, the platform can use the threat of canceled recommendations to discipline prices: sellers are prompted to lower their prices to be recommended by the algorithm. The long-run effect is less clear: if sellers do not view themselves as competitive, they may be deterred from entering the platform.

We will build the possibility of canceled recommendations into our structural model below. For now, we provide some suggestive descriptive evidence that the platform could be using the threat of canceled recommendations to discipline prices. Before 2019, Amazon's US marketplace required that third-party sellers on its platform sell their products for a lower price than on other platforms (Kelly 2019). These mandates, known as "most-favored nations clauses" (MFN), were found to be anticompetitive by European antitrust agencies, leading Amazon to drop them in those markets by 2013.[4] Following similar threats in the US late in 2018, Amazon also dropped the MFN clauses for its US sellers by March 2019.

[4]Amazon's MFN were found to be anticompetitive in Germany (see German Competition Authority 2013 and German Competition Authority 2015) and the United Kingdom (UK Competition Authority 2013). In the market for e-books, if these MFN were allowed, non-fiction book prices would be predicted to rise by 9% (De los Santos, O'Brien, and Wildenbeest 2018a).

Nonetheless, canceled recommendations may achieve a similar effect as explicit MFN. Indeed, this very issue is currently subject to litigation in American courts (District of Columbia 2021). To demonstrate the mechanism, we combine[5] data on the price $\ell_{pt}$ of the lowest-priced offer on product $p$ at date $t$ with data on Buybox ownership for our merchants as well as information[6] on the manufacturer's suggested retail price (MSRP) $R_{pt}$. We then run the following regression:

$$\begin{aligned} \text{canceled}_{pt} \quad = \quad & \alpha_p + \beta_0 \times 1\{\ell_{pt} > R_{pt}\} + \beta_1 \times \log\left(\frac{\ell_{pt}}{\mathrm{R}_{pt}}\right) \\ & + \beta_2 \times 1\{\ell_{pt} > R_{pt}\} \log\left(\frac{\ell_{pt}}{\mathrm{R}_{pt}}\right) + \epsilon_{pt}. \end{aligned} \tag{1}$$

Here, product fixed effects $\alpha_p$ soak up cross-product variation in recommendation cancellation rates. Thus, the $\beta$ parameters measure the within-product correlation between cancellation rates and the competitiveness of the lowest-priced offer. For our intended causal interpretation — that is, high prices cause canceled recommendations — we need to rule out reverse causation. If recommendations were canceled at random, our results would still show a positive correlation as long as merchants reacted to canceled recommendations by adjusting their prices upwards. However, this is implausible. After all, the primary purpose of low prices is to attract customers. Furthermore, a causal interpretation of our results aligns with comments made by industry experts. In fact, an Amazon spokesperson clarified, "if a product is not priced competitively by a seller, we reserve the right to not feature that offer" (Gonzalez 2018).

The results from the regression for Equation (1) are displayed in Figure 2. Our parameter estimates (provided in Table 11) imply that the cancellation fraction is always increasing in the price of the cheapest offer. However, this increase is much sharper if the offer price exceeds its MSRP. Indeed, for offers priced above MSRP, a 1% increase in relative price is associated with a 0.73% higher probability that the recommendation will be canceled. In the figure, we plot the relationship between $canceled_{pt} - \hat{\alpha_p}$ (on y) and $\log(\ell_{pt}) - \log(R_{pt})$ (on x). The resulting plot confirms that cancellation is increasingly likely if price rises to more than 10% above MSRP.

## 2.3 Recommendation Descriptives

If the platform decides to recommend one of the offers, how does it choose between them? We begin by providing some descriptive statistics. Figure 3a illustrates that there is within-product price variation in the data; meanwhile, Figure 3b suggests that the within-product price rank is an important determinant of recommendation status. Indeed, the cheapest offer is recommended more than 50% of the time. However, crucially, it is *not* true that the lowest-priced offer is always

[5] We collapse our data by product and date; for each such product-date pair, we call the Buybox *suppressed* if it is suppressed for every observation on that date. Similarly, we say the Buybox is not suppressed if it is not suppressed for any observation on that date. If on some date, we observe that the Buybox was suppressed at some point and not suppressed at some other point, we encode the Buybox suppression status for that date as "missing."

[6] We obtain this additional data from Keepa, a price tracker for goods on Amazon. For more details, see Data Appendix A.3.

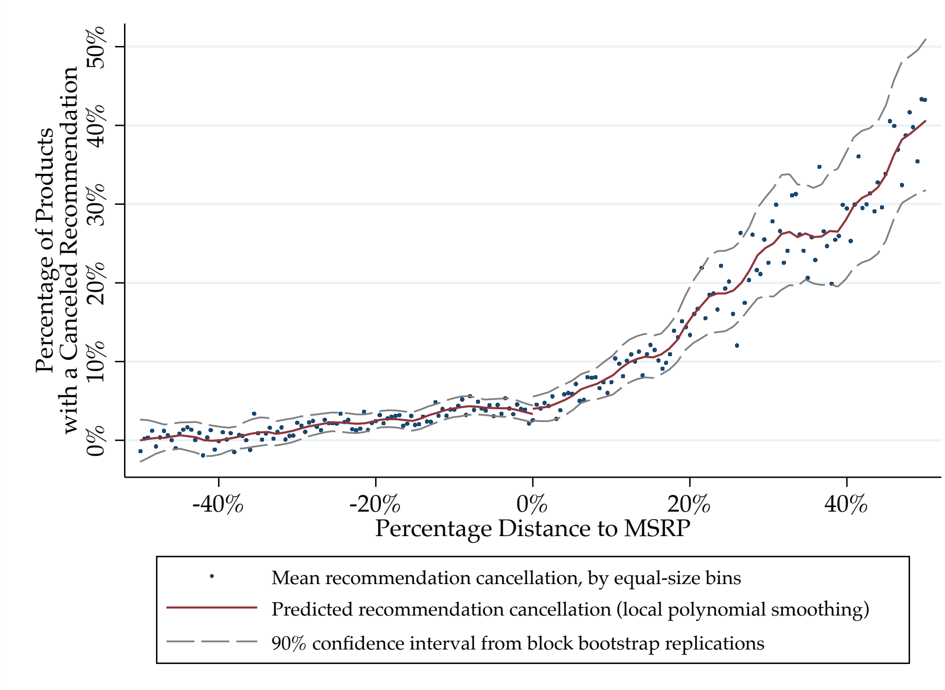



**Figure 2:** *Evidence on Canceled Recommendations (Binscatter & Lowess).*
Notes*: This figure illustrates the relationship between canceled recommendations and uncompetitively priced offers. To this purpose, we estimate (1). The resulting parameter estimates can be found in Table 11. Here, we provide model-free evidence by relating the cancellation fraction (cleaned of product FE) to the percentage distance of the lowest-priced offer's price to the MSRP. We find that an increase in price unambiguously increases the likelihood of cancellation. This effect is much more pronounced if the offer is priced above the suggested retail price for the product.*

| Variable | Coefficient |
|---|---|
| Constant | 0.075<br>[0.074,0.075] |
| $\log(\text{Price}) - \log(\text{MSRP})$ | -0.092<br>[-0.107,-0.078] |
| Recommendation Canceled? | -0.020<br>[-0.025,-0.014] |
| Product FE? | ✓ |

**Table 1:** *Canceled Recommendations Depress Sales.*
Notes*: The dependent variable is an indicator for whether there was any observed sale on a given date. This table investigates the extent to which canceled recommendations depress sales by subsetting to product-days for which we observe whether there were any sales by the lowest-priced merchant. It shows that cancelling recommendations lowers sales of the lowest-priced offer by 26.31%. We report 95% confidence intervals using standard errors clustered at the product level in square brackets.*

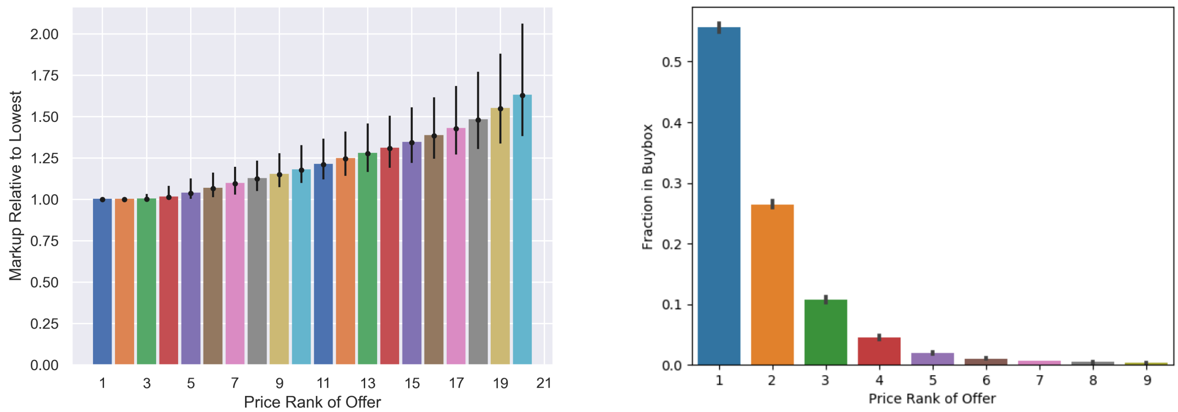
recommended. The second cheapest offer is recommended more than 25% of the time, and even more expensive offers still take an (ever decreasing) share of recommendations.

We zoom in on this phenomenon in Table 2. This table displays the fraction of observations that satisfy various criteria. For instance, we note that while the recommended offer is cheapest in 51.24% of cases, offers that are fulfilled by Amazon (FBA) are recommended in an astounding 96.46% of cases. We will see below that this is partially because FBA offers ship a lot faster than other offers. This shipping speed advantage also accrues to Amazon. While Amazon's own offers are only recommended in 5% of observations, this is mostly because we observe more data on products without Amazon offers. Once we condition on products featuring an offer by Amazon (see the the rightmost column of Table 2), we see that the platform recommends its own offer 43.85% of the time. Even among products with an Amazon offer, the same relationship between price and Buybox market share persists. We show this in Figure 7 in the Appendix.

To disentangle the various factors influencing recommendation status, we proceed by estimating a multinomial logit choice model. We relegate the technical discussion of this model to our estimation section below, including our handling of potential price endogeneity. For now, we view our results as purely descriptive. To facilitate exposition, we transform coefficients to answer questions of the following form: if the seller were to move from the 1st percentile of feedback count to the 99th percentile, by how much could she raise her price while keeping the same probability of being recommended as before? The answer is provided in Table 3: such a seller can raise her price by 2.51%, a small but statistically and economically significant amount. While the effect of increasing the percentage of positive feedback is about the same, we see that shipping time matters a lot more: after moving from the 1st to the 99th percentile of the shipping time distribution, a merchant would actually have to decrease their price by 21% to maintain her old recommendation share. In addition, there are benefits to utilizing the platform's fulfilment network, above and beyond the benefits to shipping time. Indeed, switching to using the said fulfilment network allows merchants to raise their prices by 12.56% without losing recommendation share. This benefit naturally also accrues to offers made by the platform itself. However, these offers benefit by an *additional* 9.17%, i.e. a total of 21.73%. While this additional benefit may inspire accusations of self-dealing, it is important to note that there are legitimate reasons to emphasize the platform's own offers if the customers prefer these offers. Such a preference may emerge due to consumers' worries about counterfeit goods or differences in the handling of returns.

We illustrate in Figure 4 how the recommendation benefits that accrue to an offer vary with its features. For instance, the effect of feedback count is approximately linear in log feedback count. Meanwhile, while there are large gains from achieving a positive feedback percentage above 90%, as well as further gains from increasing this percentage to around 97%, eventually these gains level off.

Finally, we can examine how the price sensitivity of the recommendation algorithm varies as the available number of offers varies. Figure 5 suggests that the recommendation algorithm becomes more price sensitive when more offers are available. This fact hints at the possibility that the platform may be wary of the effect that too stark price competition could have on entry, a theme we explore in detail below.




**(a)** *Price Multiple.*

**(b)** *Mean Recommendation Status.*

**Figure 3:** *Prices and Recommendation Status Vary Significantly with Price Rank.*
Notes*: The left figure displays the median markup of the n-th lowest-priced offer over the lowest-priced offer on each product, for various price ranks n. The right figure gives the fraction of offers at a given price rank that are recommended. Black bars in the left figure indicate interquartile ranges, while those on the right are 95% confidence intervals from mean estimation.*

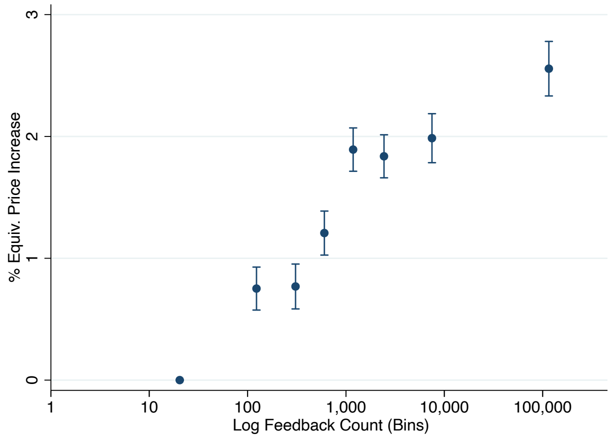
# 3 Model

In this section, we build a model of within-platform competition. Our goal is to investigate to what extent designers of online marketplaces can derive power from choosing the algorithms underlying their recommender systems. Our approach focuses on the platform's ability to guide search by influencing customers' consideration sets (Goeree 2008). While past literature has employed a similar approach to investigate platform-guided search on eBay (Dinerstein et al. 2018), our key modelling contribution lies in explicitly modelling entry. In doing so, we avoid artificially overstating the platform's power by explicitly accounting for its need to attract entrants on both sides of the market. Before delving into the details of our entry model, however, we need to discuss how consumers choose between products in the presence of a recommendation algorithm.

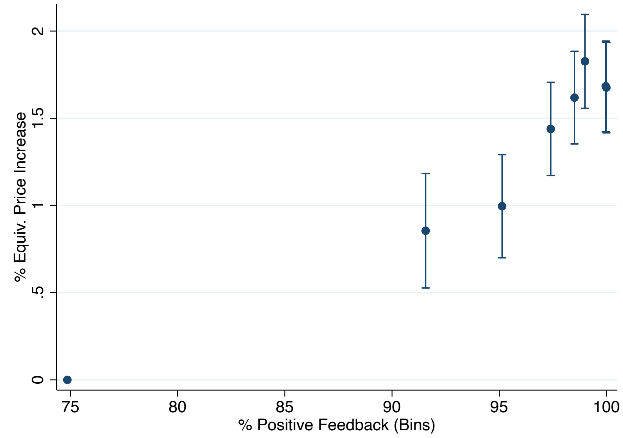
## 3.1 Consumer choice

Fix a market $t$. Where obvious, we will suppress the market subscript for ease of notation.

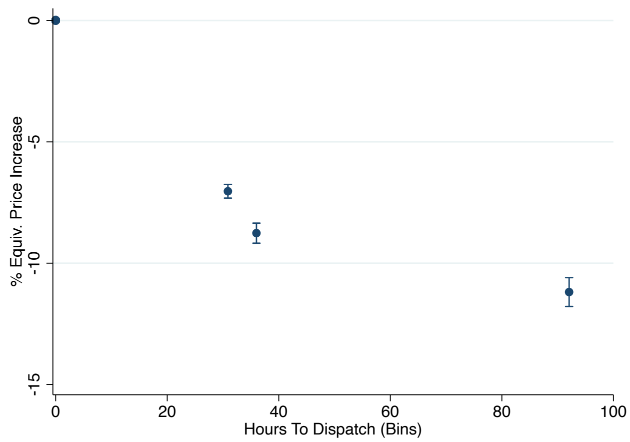
Within each market, demand follows a standard logit framework except that consumers form endogenous consideration sets as in Goeree (2008). In particular, there are two types of consumers: a fraction $\rho$ are "sophisticated." Sophisticated consumers ignore the recommendation




**(a)** *Feedback Count.*




**(b)** *Positive Feedback %.*




**(c)** *Shipping Time.*

**Figure 4:** *Non-Price Characteristics Give Recommendation (Dis-)Advantage.*

Notes*: These plots measure the extent to which a merchant could raise the price of her offer (relative to MSRP) if she were to move from the baseline value of a given feature to the value given on the x-axis. We see that both feedback count and the percentage of positive feedback matter slightly, but shipping time (measured here in hours) matters a lot more.*

| | Fraction of Observations | | |
|---|---|---|---|
| | Overall | FBA | Amazon |
| Amazon Offer Exists | 11.28% | 11.34% | 100.00% |
| FBA Offer Exists | 99.44% | 100.00% | 99.98% |
| Winner Is ... | | | |
| Lowest Priced | 51.24% | 51.20% | 45.43% |
| Second Lowest Priced | 28.31% | 28.32% | 31.84% |
| Highest Feedback Count | 17.23% | 17.14% | 8.32% |
| Highest Feedback Rating | 18.13% | 18.11% | 15.02% |
| Fastest Shipping | 19.61% | 19.60% | 23.45% |
| Lowest Priced FBA | 63.47% | 63.84% | 63.26% |
| FBA | 96.46% | 97.01% | 98.07% |
| Amazon | 4.95% | 4.98% | 43.85% |
| Lowest Fastest Shipping | 64.59% | 64.76% | 64.11% |

**Table 2:** *Determinants of Recommendation Status.*
Notes*: This table provides the fraction of recommendation observations satisfying the criteria listed in the first column. The second column is based on the entire set of observations, the third column employs only observations for which there is an offer that is fulfilled by Amazon and the fourth column uses only observations for which there is an offer by Amazon itself.*

and evaluate all available options $\mathcal{J} \cup \{0\}$. The remaining $1-\rho$ consumers are "unsophisticated." These consumers only consider the recommended offer $j^r \in \mathcal{J}$ and the outside option. Thus, the recommendation algorithm will essentially be allowed to "choose" which offer an unsophisticated consumer evaluates. On Amazon Marketplace, consumers that do not explicitly click through to the offer listing are indeed never informed of the availability or characteristics of non-recommended offers. This feature of our setting supports the applicability of the consideration set modeling framework.

In each market $t$, both types of consumers have preferences over the alternatives available to them. Their mean utilities for alternative $j$ depend on its characteristics $\mathbf{x}_{jt}$, price $p_{jt}$ and unobserved quality $\xi_{jt}$:

$$\delta_{jt} = \mathbf{x}_{jt}'\beta - \alpha_t p_{jt} + \xi_{jt}.$$

This formulation allows the price coefficient $\alpha_t$ to depend on the market $t$; we will fill in the details of this dependence when we discuss estimation. As in typical demand estimation, only differences in mean utilities with respect to a "reference option" are identified. Hence, we normalize the mean utility of the outside option to zero: $\delta_{0t} = 0$. Mean utilities on each offer are combined in the usual fashion with a Type-1 extreme value shock $\epsilon_{ijt}$, forming indirect utilities

$$v_{ijt} = \delta_{jt} + \epsilon_{ijt}.$$

Each consumer then chooses the option in his consideration set that maximizes his utility. Therefore, given a recommended offer $j^r$, the probability that a consumer chooses product $j$ in market $t$

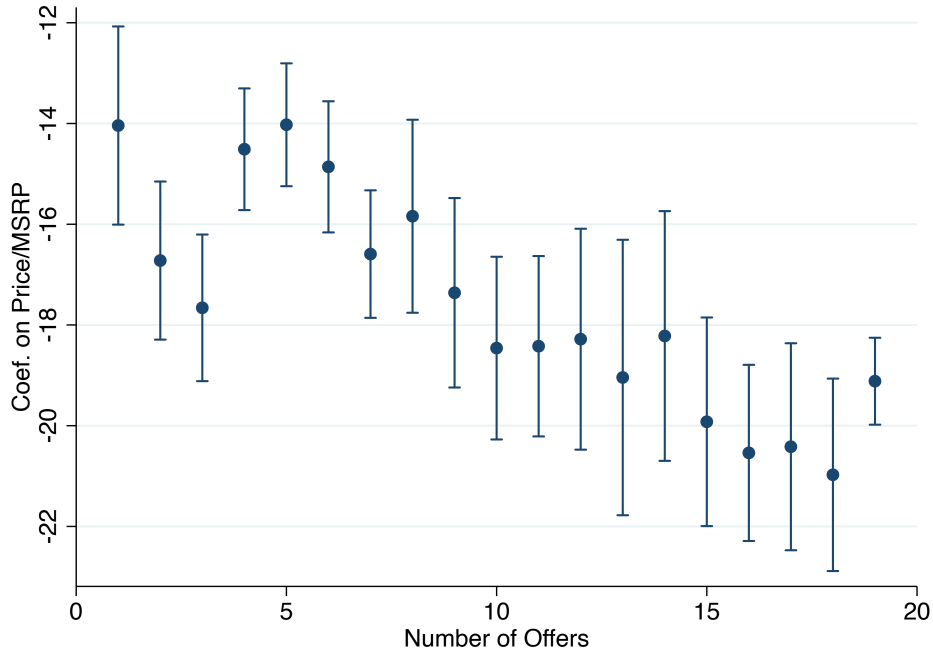



**Figure 5:** *Recommendations Are More Price Sensitive When There Are More Offers.*
Notes*: We illustrate how the coefficient on price divided by suggested retail price varies as the number of offers varies.*

| | 1st | 99th | Effect |
|---|---|---|---|
| Is FBA? | 0 | 1 | 12.56% [11.74%,13.37%] |
| Is Amazon? | 0 | 1 | 9.17% [8.84%,9.51%] |
| Feedback Count | 1 | 179,594 | 2.51% [2.29%,2.73%] |
| Pos. Feedback % | 50 | 100 | 1.61% [1.36%,1.86%] |
| Shipping Time | 0 | 192 | -20.86% [-22.02%,-19.70%] |

**Table 3:** *Recommendation Equivalent Price Effects.*
Notes*: This table provides the amount by which a merchant would have to increase price (as % relative to MSRP) in order to keep his recommendation share constant after moving from the 1st percentile to the 99th percentile of the covariate given in the left-most column.*

is

$$d_{jt}(j^r) \quad = \quad \rho \times \frac{\exp(\delta_{jt})}{1+\sum_{j\in\mathcal{J}_t}\exp(\delta_{jt})} + (1-\rho)\times \mathbf{1}\{j=j^r_t\}\times\frac{\exp(\delta_{jt})}{1+\exp(\delta_{jt})}.$$

### 3.2 Recommendation Algorithm

A *recommendation algorithm* maps characteristics of alternatives to a recommended alternative $j^r \in \mathcal{J}$ or, if no recommendation is given, the null recommendation $j^r = \{0\}$. While, in general, the recommendation may vary by consumer, this is not the case in our empirical application. Hence, we model the recommender system as solving exactly one discrete choice problem for each market.[7] The mean utility $\delta^r_{jt}$ of each inside alternative includes observable characteristics $\mathbf{x}_{jt}$, price $p_{jt}$ as well as the econometrician-unobservable quality $\xi_{jt}$:

$$\delta^r_{jt} = \mathbf{x}'_{jt}\beta^r_t - \alpha^r_t p_{jt} + \xi^r_{jt}.$$

As with demand, we allow the price coefficient $\alpha^r_t$ to depend on market characteristics and normalize the utility of the outside option, $\delta^r_0 = 0$.

Whether due to deliberate randomization or mistakes in evaluation, we will assume that recommendation decisions are based on a shocked version $v^r_{jt}$ of this mean utility.[8] In particular, we will assume a nested logit structure. Thus, we partition the set of products as $\mathcal{J}_t \cup \{0\}$ and let $g = 1$ be the index of the nest of inside options. The recommender utility of product $j$ is then

$$v^r_{jt} = \delta^r_{jt} + \zeta^r_{gt} + (1-\lambda)\epsilon^r_{jt},$$

where $\epsilon^r_{jt}$ is distributed i.i.d. Type-1 Extreme Valued, and $\zeta^r_{gt}$ is common to all options in the same nest. By Cardell (1997), the distribution of $\zeta^r_{gt}$ can be specified in such a way that $\zeta^r_{gt} + (1-\lambda)\epsilon^r_{jt}$ has a Generalized Extreme Value distribution. Finally, alternative $j$ is recommended in market $t$ if and only if it has the highest recommender utility among all options. This happens with probability

$$r_{jt} = \frac{\left[\sum_{k\in\mathcal{J}_t}\exp(\delta^r_{kt}/\lambda)\right]^{\lambda}}{1+\left[\sum_{k\in\mathcal{J}_t}\exp(\delta^r_{kt}/\lambda)\right]^{\lambda}} \times \frac{\exp(\delta^r_{jt}/\lambda)}{\sum_{j\in\mathcal{J}_t}\exp(\delta^r_{kt}/\lambda)}.$$

Combining this with our previous result on demand, this yields market shares

$$s_{jt} = \rho\times\frac{\exp(\delta_{jt})}{1+\sum_{j\in\mathcal{J}_t}\exp(\delta_{jt})} + (1-\rho)\times r_{jt}\times\frac{\exp(\delta_{jt})}{1+\exp(\delta_{jt})}.$$

[7] Formulating the platform recommendation problem in this manner makes it compatible with our discrete-choice model of demand. The main distinction between modeling recommendation and consumer choices lies in the fact that the platform may only make *one* discrete choice per market. In contrast, there are usually multiple consumers, each making their own choice in each market in typical discrete choice demand models.

[8] We provide evidence for randomization on the Amazon platform in Appendix B.

### 3.3 Firm Choices

We model entry and competition across a set of markets $\mathcal{T}$. Each market $t$ is associated with a fixed cost $F_t$ as well as a set of potential entrants $j \in \mathcal{N}_t$ who draw a type $\omega_j \sim G(\cdot)$ and play an entry-and-pricing game. Each firm $j$ chooses a pair $(\chi_j, p_j)$ consisting of an entry and a pricing strategy. While $\chi_j : \mathcal{I}_j \to \{0,1\}$ maps a player's pre-entry information set $\mathcal{I}_j$ into her entry decision, $p_j : \mathcal{I}'_j \to \mathbb{R}^+$ maps a player's post-entry information set $\mathcal{I}'$ into the price she will charge for her product. Allowing for the possibility that merchants receive only a share $\phi$ of the revenue they generate (as in our application), payoffs are given by

$$\pi_j(\boldsymbol{\omega}, \boldsymbol{\chi}, \mathbf{p}) = \sum_{t \in \{\tilde{t} | j \in \mathcal{N}_{\tilde{t}}\}} \chi_{jt} \times \left[ \left( \phi p_{jt} - C_t \left( s_{jt}(\boldsymbol{\omega}, \mathbf{p}), \omega_j \right) \right) s_{jt}(\boldsymbol{\omega}, \mathbf{p}) - F_t \right],$$

where $s_{jt}$ maps the vectors of types and prices into demand. Our solution concept is Bayes-Nash Equilibrium. We proceed by stating several assumptions that are necessary to keep our model tractable.

**Assumption 1** (Separable Markets)**.** *Each firm enters in at most 1 market: $\mathcal{N}_t \cap \mathcal{N}_s = \emptyset$ for $t \neq s$.*

**Assumption 2** (Unit Production Costs)**.** $C_t(\cdot, \omega_j) = c_j \in \mathbb{R}^+$.

**Assumption 3** (Full Information Pricing)**.** *After having entered, each seller knows the identities, cost and quality draws of its opponents when playing the pricing game: $\mathcal{I}' = (\boldsymbol{\omega}, F)$.*

**Assumption 4** (Equilibrium Selection)**.** *When there are multiple pricing equilibria, firms play the pricing equilibrium in which the firm with the lowest-cost offer attains the highest recommendation share.*

**Assumption 5** (Blind Entry)**.** *Before entry, each seller only knows its own unit production cost draw and the (common) fixed cost of entry: $\mathcal{I} = (c_j, F)$.*

**Assumption 6** (Symmetric Entry)**.** $\chi^*_j(c, F) = \chi^*_k(c, F)$ *for all* $j, k \in \mathcal{N}$ *and* $c, F \in \mathbb{R}^+$.

The separable markets assumption allows us to solve for equilibrium market-by-market, easing the considerable computational burden of our model. Its validity hinges on the role cross-market cannibalization plays in firms' pricing strategies. In our empirical application, firms are small and rarely specialize in a related set of products (i.e., there is no monopoly of socks on our platform). Thus cross-market cannibalization considerations presumably are only a secondary concern for the sellers we observe.

Assuming unit production costs is common in the literature and particularly plausible in our context as sellers are retailers (as opposed to producers). Full information pricing is standard. Jointly, these assumptions imply that successful entrants play a Bertrand-Nash pricing game (Anderson, De Palma, and Thisse 1992). In our simulations, we find that at the pricing stage, equilibria always exist but are not necessarily unique. This is because "quantities" are determined by the induced market share $s_j(\cdot)$, which includes the effect of the recommendation on consumer choice. Since the algorithm's behavior is known to sellers, each will internalize its effect on demand when setting prices. Indeed, we have

$$\frac{\partial s_j}{\partial p_j} = \rho \frac{\partial s_j^{sophisticated}}{\partial p_j} + (1-\rho)\left[\frac{\partial r_j}{\partial p_j} s_j^{unsophisticated} + r_j \frac{\partial s_j^{unsophisticated}}{\partial p_j}\right].$$

Thus, through influencing $\frac{\partial r_j}{\partial p_j}$, the recommendation system partially determines the price elasticity that sellers face — this is where the platform's power to guide consumer search manifests in our model. However, the added complication of recommendations can create multiplicity issues. Intuitively, firms will attempt to differentiate by pursuing *either* sophisticated *or* unsophisticated consumers, and different permutations of these roles can create multiple equilibria. Where there are such multiplicities, we employ Assumption 4 to select an equilibrium.[9]

The last two assumptions provide us with unique, computationally tractable equilibria at the entry stage. In each market, a set of potential entrants (sellers) $\mathcal{N}$ face a fixed cost of entry $F$. Each putative entrant independently decides whether to enter after observing information set $\mathcal{I}_j$ which the blind entry assumption restricts to $\mathcal{I}_j = \{c_j, F\}$. This strong assumption consists of two parts. Firstly, in the spirit of Roberts and Sweeting (2013), we deprive sellers of knowledge about offers *other* than their own. This is necessary to avoid the classic equilibrium multiplicity problems that emerge in games of strategic substitutes (e.g., Tamer 2003). It is also innocuous in our context because most sellers in our sample are small, anonymous merchants who are mostly unaware of their competitors before entry. Even when they know each other's identity, turnover is high. Therefore, merchant entry is best modeled with firms expecting to play against random draws from the distribution of potential opponents.

The second restriction has more bite: before entry, the only information each firm has about its type is its cost. This is plausible to the extent that the recommendation algorithm is opaque: entrants are uncertain, *ex-ante*, about the non-cost factors that algorithm and consumers value, and only learn about them after entry (e.g. Jovanovic 1982). Prima facie, it would be straightforward to dispense with this assumption from a game-theoretic view. That is, we could allow entry decisions to depend on private information about vertical characteristics. However, we require the assumption to keep the computation tractable.[10] Nevertheless, we note that the assumption could influence results as it rules out a possible role of the recommendation algorithm as selecting entrants on quality. To the extent that this is the case, our analysis below will overstate the welfare loss due to recommender-inspired exit.

Each firm will enter if and only if its expected profits from entering exceed fixed cost $F$. Given our assumptions, this expectation depends only on a firm's own unit cost draw. As

[9] To assess the robustness of this assumption, we also estimated stage games and selected equilibria using alternative rules. These are: (1) a random seller has the lowest price, and (2) the seller with the highest consumer demand attractiveness has the lowest price. In either case, our results remain qualitatively unchanged.

[10] The entry game is solved many times in an inner loop when we search for parameters in our estimation procedure. We are currently working on relaxing this assumption to allow for simultaneous selection on cost and demand attractiveness.

its expected profits are declining in its own unit cost $c_j$, each firm best responds in cutoff strategies: $\chi_j(c_j) = \mathbf{1}\{c_j < c_j^*\}$. Thus, an equilibrium can be characterized by a vector $(c_1^*, \ldots, c_n^*)$. Assumption 6 further restricts attention to symmetric equilibria. Thus, all (ex-ante identical) firms share the same (ex-ante) cutoff $c^*$. In Appendix C.1, we show

**Proposition 1.** *The entry game has a unique symmetric equilibrium in cutoff strategies.*

This unique cutoff $c^*$ exactly balances expected gross profits against fixed costs if other firms enter according to our putative cutoff rule. That is, each firm enters if and only if it draws a unit purchase cost weakly below $c^*$. Formally, $c^*$ solves

$$\mathbb{E}_{\mathbf{q},c_{-j}}[\pi_j(c^*, q_j; c_{-j}, q_{-j}, \chi^*_{-j})] = F \text{ where } \chi_k^*(c) = 1\{c_k < c^*\} \text{ for all } k \neq j.$$

(In a slight abuse of notation, we here use $\pi$ as the downstream profits gross of fixed costs.) This equation is what we exploit in our estimation to implement our numerical search for equilibria as discussed in Appendix E.2.

# 4 Estimation

Since we have access to data on more than 200,000 products, an estimation procedure run on the full data set will be computationally intensive. By restricting the estimation sample, we can keep the products as comparable as possible. Thus, we focus on all 11,148 products in the Fashion category[11] on Amazon Marketplace for which we have data. We define a market $t$ to be a "product page"-day pair, $(p, \tau)$. In each market, the alternatives $\mathcal{J}_t$ are the various offers on the same product page. This market definition allows us to sidestep having to estimate potentially complex preferences over product characteristics: all such characteristics will be common amongst offers for the same product, so they cancel out in the discrete choice problem. Nevertheless, the various offers may still differ in characteristics such as shipping speed.

Traditionally, distinct "markets" are either repeated observations of the same market at different times (e.g. Berry, Levinsohn, and Pakes 1995) or geographically distinct markets in which consumers are offered (nearly) the same alternatives (e.g. Nevo 2001). By contrast, in our data, the kind of alternatives offered to customers varies product page by product page. It seems natural to suspect that preferences over offer characteristics might depend on what the product is. In theory, we could allow preferences to freely vary by product page. However, in practice, this approach is underpowered. While we observe more than 100,000 sales, these are distributed across more than 10,000 products. To keep estimation tractable, we impose a functional form assumption: when comparing various offers for a given product, consumers evaluate these offers relative to the manufacturer's suggested retail price (MSRP) of said product. As an example, consider a pen and laptop, costing \$10 and \$1,000 respectively. For the pen, it is plausible that a

[11] We choose this category as it is the category for which we have the most extensive sales data and the limiting factor in our estimation will be power to identify the determinants of demand.

difference of \$0.50 may sway a consumer's choice from one offer to another. Yet, for the laptop, the relevant price difference is typically closer to \$50, not \$0.50. Formally speaking, if $R_p$ is the MSRP of product $p$, we assume consumers evaluate offers on the basis of the ratio between the product's price and $R_p$, i.e.

$$\alpha_{p,\tau} = \frac{\overline{\alpha}}{R_p}, \quad \text{and} \quad \alpha^r_{p,\tau} = \frac{\overline{\alpha}^r}{R_p}.$$

As we show in Appendix D.6, this functional form assumption fits our data well. Furthermore, it is supported by our conversations with industry experts.

### 4.1 The Recommendation Algorithm

We estimate the parameters of the recommendation algorithm by maximum likelihood (MLE) and address the potential endogeneity of price using offer fixed-effects. The details of introducing these fixed-effects into a maximum likelihood procedure without causing an incidental parameters problem are relegated to Appendix D.4. Essentially, we apply Chamberlain's (1980) conditional logit approach to the conditional logit model itself (see also Rasch 1960, 1961). This approach yields an estimator that exclusively exploits within-offer price variation to identify $\alpha$. The identifying assumption underlying our estimation strategy is

**Assumption 7.** *Offer qualities are time-invariant, i.e.* $(\xi^r_{jt}, \xi_{jt}) \equiv (\xi^r_j, \xi_j)$.

Our estimation strategy substantially diverges from prior literature for two reasons. Firstly, we have access to high-frequency, extremely disaggregate data. Thus, if we were to recover mean utilities by log-transforming the ratio of market shares as in Berry (1994), our estimates would suffer from severe zero-share bias (Gandhi, Lu, and Shi 2017; Quan and Williams 2018). Secondly, our data comes from a dynamic marketplace characterized by frequent price changes and high merchant turnover. These features suggest the time dimension contains more identifying variation than common in the literature. Furthermore, especially given the delegation of pricing to algorithms reacting to competitors' prices[12], this variation is plausibly exogenous. The usual strategy to address zero-share bias, aggregation, would thus smooth precisely over the most informative variation. Worse, in the absence of a plausible instrument for price, it would introduce measurement error and attenuate the estimated coefficients.

To avoid the issues, we estimate our parameters via MLE. This, however, requires us to address potential endogeneity concerns. Since the estimating equation is non-linear, we cannot pursue an instrumental variable approach. The usual worry is that the unobserved quality of an alternative may be correlated with its price. However, recall that in our context, a market is a product page. Thus, unobserved *product* quality is the same between offers on the same product and hence cancels. Instead, $\xi^r_j$ refers to the unobserved *offer* quality. One may argue that *unobserved* offer quality is unlikely to matter much because we observe the key determinants of offer quality. For instance, we see an offer's time to dispatch, whether the offer is fulfilled by Amazon's own

[12] Algorithms that react to demand conditions are not offered by the repricing company that is our data source.

logistics operation, whether the offer is listed by Amazon itself, as well as other measures of seller quality. However, consumers may still use the seller's name to draw inferences about seller quality that are not visible to the econometrician.

Nevertheless, our data on observable offer characteristics motivate a natural strategy to deal with endogeneity. In Table 4, we exhibit the $R^2$ from regressing various observable offer characteristics on offer fixed effects. These $R^2$ are all very close to one, thus suggesting that essentially all variation in offer characteristics comes from cross-sectional rather than temporal variation. It is plausible, then, that unobservable quality is time-invariant as well, i.e. we believe our identifying Assumption 7 holds.

| **Characteristic** | **$R^2$ on offer FE** |
|---|---|
| Shipping | 0.95 |
| Feedback | 1.00 |
| FBA | 0.97 |
| Amazon | 1.00 |

**Table 4:** *Observable Offer Quality is Time Invariant.*
Notes: *This table exhibits the $R^2$'s obtained by regressing an observed product characteristic (left) on estimated offer fixed effects. As the $R^2$ is above 0.95 in all cases, we conclude that observable offer quality is time invariant.*

We exhibit the results of our estimation of the recommendation algorithm in Table 5. Specifications (1) and (3) ignore offer quality, i.e. they implicitly assume $\xi_j^r \equiv 0$. Estimation of (1) and (3) is thus by MLE. By contrast, specifications (2) and (4) treat offer quality as a random effect $\xi_j^r \sim N(0, \sigma^2)$. The estimated standard deviation of this random effect, $\sigma$, is reported as "RE" in the table. We estimate models (2) and (4) by maximum simulated likelihood. Finally, specification (5) employs offer fixed-effects; hence, it is robust to arbitrary correlation between offer quality and price under Assumption 7.

We emphasize a few findings.[13] First of all, comparing columns (4) to (5), we find no evidence of endogeneity beyond that accounted for by observed quality covariates. The implied price coefficient is about -8 in both specification (4) with observed quality covariates and specification (5) which controls for offer fixed-effects. Second, the implied price elasticity of the recommendation algorithm is extremely high, at about -19. This finding is corroborated by our discussions with sellers on the platform. Even more intriguingly, on its seller interface, Amazon explicitly tells the seller which of his "listings [...] are priced not more than 5% above the Buybox." This statement is consistent with our finding that offers priced more than 5% above the cheapest offer receive almost no recommendations. Thirdly, as $\hat{\lambda} < 1$, we find strong evidence that Amazon considers offers more substitutable with each other than with the outside option. This is, of course, what one would expect given that choices of the outside option do not yield any intermediation fees. Finally, slow shippers are penalized: shipping one day more slowly is equivalent to having an offer that is 9% more expensive. Similarly, Amazon's offers appear to have an (additional) advantage

[13] Note that our results here are expected to vary slightly from our discussion in the descriptive analysis section as the sample has changed.

equivalent to a 9% price discount. Does this advantage reflect a quality difference, or is it merely self-dealing? Below, we will see that our counterfactual results appear to support the quality hypothesis.

### 4.2 Consumer Choice

We face the same challenges of endogeneity and high-frequency data when estimating consumer choice as in the last section. However, it becomes even more critical not to smooth over short-run temporal variation because of how the fraction of unsophisticated consumers, $\rho$, is identified. Intuitively, $\rho$ is pinned down by the observed covariance between sales and recommendation status. As the latter is subject to high-frequency variation, we want to avoid smoothing over this variation by aggregating over time. Thus, as before, we proceed by MLE and employ a fixed-effects strategy to address endogeneity.

We face an additional complication in estimating demand: *partial observability*. For each market, there is precisely one alternative for which we observe how often consumers choose it. Conditional on knowing the market size, partial observability raises no identification challenges (Matzkin 2007) and is addressed with a straightforward modification to the likelihood function. However, partial observability means we cannot condition out offer-level fixed effects: essentially when one only observes a single merchant's sales, it is impossible to condition on pairs of product-days such that a sale goes to one firm on the first and another firm on the second. In Appendix D.5, we show how we can exploit the high-frequency nature of our data to condition out fixed-effects nonetheless. To do so, we focus on pairs of days between which all except the observed merchant's prices and other offer characteristics stay fixed. When this is the case, progress can be made by comparing the sales of our observed merchant on days where his price is high to his sales on days where his price is low.

To estimate the market size, we employ additional data on sales ranks in a procedure described in detail in Appendix D.1. Since their introduction into the literature in Chevalier and Mayzlin (2006), sales ranks on Amazon have frequently been used as a proxy for total product-level sales (e.g. Reimers and Waldfogel 2019, De los Santos, O'Brien, and Wildenbeest 2018b). In this vein, we estimate a linear relationship between observed log sales and log sales ranks for the sample of products with just one seller. This allows us to predict total sales for each market. We then translate this figure to a market size by dividing by the percentage of sessions on the e-commerce platform that result in sales: 12.3% (Berthene 2017).

The results of our consumer choice estimation can be found in Table 6. Our preferred specification is (9), which incorporates our consideration set model (i.e., does not restrict $\rho$) and observable quality covariates. We find that $1-\rho \approx 26\%$ of consumers are "naive" in that they only consider the recommended offer and the outside option. It is interesting to note that this estimate is similar to the fraction of users that complete Amazon purchases in three minutes or less: 28% (Bezos 2021). The overall price elasticity of demand is around -4. By comparing the price elasticity holding recommendations fixed ($\text{Elasticity}_1$) to the one where recommendations

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Price | | | | | -8.44<br>[-10.67,-6.21] |
| Inside | 6.61<br>[8.17, 5.04] | 7.88<br>[9.84, 5.92] | 6.49<br>[8.09, 4.88] | 7.63<br>[9.64, 5.63] | |
| Price / MSRP | -4.02<br>[-5.86, -2.18] | -5.35<br>[-7.61, -3.09] | -3.93<br>[-5.90, -1.96] | -5.07<br>[-7.35, -2.80] | |
| Shipping | | | -0.30<br>[-0.56, -0.05] | -0.46<br>[-0.55, -0.38] | |
| Feedback | | | -0.00<br>[-0.01, 0.00] | -0.01<br>[-0.02, -0.01] | |
| FBA | | | 0.07<br>[-0.29, 0.43] | -0.01<br>[-0.07, 0.04] | |
| Amazon | | | 0.35<br>[0.29, 0.40] | 0.46<br>[0.37, 0.54] | |
| RE | | 0.57<br>[0.46, 0.68] | | 0.42<br>[0.34, 0.51] | |
| $\lambda$ | 0.30<br>[0.23, 0.38] | 0.38<br>[0.31, 0.45] | 0.22<br>[0.15, 0.30] | 0.31<br>[0.24, 0.38] | |
| Offer FE? | ✗ | ✗ | ✗ | ✗ | ✓ |
| Impl. Inside Price | -6.44 | -6.77 | -8.55 | -8.00 | -8.44 |
| Impl. Elasticity | -15.22 | -16.03 | -20.17 | -18.90 | n/a |

**Table 5:** *Recommender System Estimates.*

Notes*: This table contains the results from MLE estimation of the recommender system using 1,345,563 observations on recommendation status. The reported coefficients measure the effect of price (in 10 USD), the mean utility of the inside option (relative to the outside option), the effect of an extra day until dispatch, the effect of a percentage increase in the number of seller reviews, the effect of an offer being fulfilled by Amazon and the effect of an offer being listed by Amazon itself. The number in row "RE" represents the standard deviation of the estimated random effect.* $\lambda$ *is the nesting coefficient. Finally, "Elasticity" refers to the average price elasticity of the recommender system.*

are allowed to vary ($\text{Elasticity}_2$), we conclude that the recommendation algorithm intensifies price competition by increasing the former elasticity by 25%.

Regarding non-price characteristics, consumers prefer sellers with a large amount of experience as a 1% increase in feedback count is equivalent to a 2% decrease in price. However, we find that while our preferred model indicates a dislike of additional shipping time that is of the same order as that estimated in for the recommendation algorithm, the estimated effect is not statistically significant. Similarly, the effect of being fulfilled by Amazon's logistics network is estimated very imprecisely. Like the recommender system, however, consumers also have a strong and statistically significant preference for offers by Amazon. We caution, however, that this effect is estimated exclusively from demand data on offers not listed by Amazon. Put differently, we infer the demand on Amazon offers from observing the "shadow" that Amazon casts in terms of reduced sales on our third-party merchants.

We now move on to discussing the remaining non-preferred models exhibited in columns (1) to (8) of Table 6. These models act as robustness checks and help us build intuition for the data generating process. The first five columns contain estimates from models that restrict $\rho = 1$, i.e., assuming all consumers are sophisticated. These estimates allow us to compare specifications with offer-level fixed effects (5) to those without (1–4). We find that we underestimate price sensitivity when we do not model unobserved quality (1, 3). Meanwhile, modeling unobserved quality as a random effect (2, 4) yields estimates of the price coefficient comparable to those under fixed effects. Hence, we conclude that endogeneity is only a minor concern.

The second set of columns (5–9) drops the restriction that $\rho = 1$. When $\rho$ is freely estimated, all our models agree that $\rho < 1$, i.e., a significant fraction of unsophisticated consumers only considers the option recommended to them. The magnitude of $\rho$ depends on whether or not we include random effects in the model. To the extent that we do not observe all changes in recommendation status, random effects could soak up some of the tendency of certain offers to be recommended more often than others. Hence, the models with random effects may overestimate the fraction of sophisticated consumers. However, there may be large quality differences between offers that, while unobservable to the econometrician, are observed by the recommender system. In this case, (6) and (8) may overstate the influence of the recommendation. Therefore, we err on the side of caution and take (9) to be our preferred specification.

### 4.3 Estimating wholesale and fixed costs

In our model, merchant entry and profits depend on their wholesale and fixed costs $(c_j, F)$ respectively. Since entry is selective on marginal costs, we cannot invert prices to costs without biasing our mean cost estimates downwards. Instead, we will parametrize the wholesale and fixed cost distributions and estimate our model via the Simulated Method of Moments (SMM) (McFadden 1989; Pakes and Pollard 1989). Since this procedure is computationally complex, we believe that our main technical contribution to the literature is to render it feasible.

For illustration, fix a market $p$. To ensure that there can be no negative realizations of wholesale

| | No Recommender Sys. | | | | | Recommender Sys. | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Price / MSRP | -1.59<br>[-3.31, 0.13] | -2.56<br>[-3.84, -1.28] | -1.70<br>[-3.50, 0.10] | -3.12<br>[-4.44, -1.80] | -3.72<br>[-5.26,-2.17] | -1.42<br>[-3.15, 0.31] | -2.47<br>[-3.66, -1.27] | -1.47<br>[-3.24, 0.30] | -2.42<br>[-4.28, -0.57] |
| Inside | -1.24<br>[-2.96, 0.49] | 0.22<br>[-1.36, 1.81] | -1.62<br>[-7.12, 3.89] | 0.96<br>[-1.63, 3.55] | | -1.01<br>[-2.71, 0.69] | 0.26<br>[-0.81, 1.33] | -1.25<br>[-6.10, 3.59] | 0.20<br>[-5.48, 5.88] |
| Shipping | | | 0.23<br>[-0.47, 0.93] | -0.44<br>[-0.93, 0.05] | | | | 0.15<br>[-0.45, 0.75] | -0.30<br>[-0.76, 0.17] |
| Feedback | | | 0.05<br>[0.01, 0.08] | 0.08<br>[0.04, 0.13] | | | | 0.06<br>[0.02, 0.10] | 0.05<br>[0.01, 0.09] |
| FBA | | | 0.60<br>[-4.33, 5.53] | -0.02<br>[-1.96, 1.92] | | | | 0.32<br>[-3.74, 4.39] | -0.11<br>[-2.69, 2.48] |
| Amazon | | | 3.99<br>[3.60, 4.37] | 3.16<br>[2.90, 3.42] | | | | 4.20<br>[3.64, 4.76] | 2.88<br>[2.17, 3.60] |
| RE | | 1.93<br>[1.81, 2.04] | | 1.96<br>[1.90, 2.02] | | | 2.04<br>[1.95, 2.14] | | 2.03<br>[1.92, 2.14] |
| $\rho$ | 1.0<br>[constr.] | 1.0<br>[constr.] | 1.0<br>[constr.] | 1.0<br>[constr.] | | 0.48<br>[0.42, 0.55] | 0.74<br>[0.70, 0.78] | 0.52<br>[0.45, 0.59] | 0.74<br>[0.70, 0.78] |
| Offer FE? | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ |
| Elasticity$_1$ | -2.00 | -3.11 | -2.13 | -3.76 | n/a | -1.77 | -2.95 | -1.83 | -2.90 |
| Elasticity$_2$ | | | | | | -3.10 | -3.87 | -3.09 | -3.80 |

**Table 6:** *Demand Estimation Results.*

Notes*: This table was produced using MLE with S=25 simulations on N=11,148 products with an average of 120.7 observations and 13.6 sales per product. The reported coefficients measure the effect of price (in 10 USD), the mean utility of the inside option (relative to the outside option), the effect of an extra day of time until dispatch, the effect of a percentage increase in number of feedback and the effect of an offer being fulfilled by Amazon. $\rho$ is the fraction of consumers that are sophisticated. Elasticity$_1$ refers to the average price elasticity of demand assuming that price does not influence consideration sets and Elasticity$_2$ refers to the same using the earlier estimates according to which it does. The underlying Buybox elasticity is -18.90.*

costs, we will parameterize them as lognormal, i.e.

**Assumption 8** (Unit Costs DGP)**.** $c_{jp} \sim LogN(\theta_0^c + x_p'\theta_x^c,\ \theta_\sigma^c)$.

Here, $\theta^c = (\theta_0^c, \theta_x^c, \theta_\sigma^c)$ are parameters to be estimated. Due to computational constraints, in practice we include only the suggested retail price of the product in $x_p$.

Similarly, we also parametrize fixed costs as being lognormally distributed:

**Assumption 9** (Fixed Costs DGP)**.** $F_p \sim LogN(\theta_0^F + x_p'\theta_x^F, \theta_\sigma^F)$.

Here, $\theta^F = (\theta_0^F, \theta_x^F, \theta_\sigma^F)$ are parameters to be estimated. Estimating fixed costs is less costly, so we can allow the mean of this distribution to vary more flexibly by including more variables in $x_p$. Thus, we include both the suggested retail price as well as the market size.

Identification is standard. The entrant distribution allows us to draw conclusions about the fixed cost parameters $\theta^F$: the mean number of entrants is informative about $\theta_0^F$; its covariance with market characteristics tells us about $\theta_x^F$; finally, its variance speaks to $\theta_\sigma^F$. Similarly, the price distribution contains all the required information about wholesale cost parameters: the mean price across markets provides a moment that speaks to $\theta_0^c$; its covariance with the suggested retail price gives us $\theta_x^c$; and the within-market price variance tells us about $\theta_\sigma^c$.

Estimation is computationally challenging. To evaluate the moments for given parameters $\theta$, we need to solve the entry game for each of 11,148 markets. However, to solve a single entry game, we need to find its associated wholesale cost cutoff $c_p^*$. Doing so requires us to compute equilibrium profits under many candidate wholesale cost values. Finally, we need to evaluate many candidate parameters $\theta$ during our outer GMM optimization procedure.

Therefore, we use several tweaks to keep estimation computationally feasible. To begin with, we speed up computation of the pricing game equilibrium by chaining distinct fixed-point iterations. While it is conventional to use the fixed-point algorithm of Berry, Levinsohn, and Pakes (1995, henceforth BLP), we employ instead the $\zeta$-markup equation of Morrow and Skerlos (2011, henceforth MS), which has stronger local convergence properties[14] (Conlon and Gortmaker 2020). Moving up to the entry game, we adopt an endogenous grid approach to reduce the number of stage game evaluations needed to find the wholesale cost entry cutoffs. Moreover, since expected gross profits are discontinuous in the number of entrants, we smooth over these discontinuities via importance sampling (Ackerberg 2009). Finally, to speed up the outer loop, we concentrate out the wholesale cost parameters and employ a modern derivative-free[15] optimization algorithm (Cartis et al. 2019, henceforth DFO-LS). This new solver separately models each moment's response to each parameter, resulting in performance superior to the classic simplex algorithm (Nelder and Mead 1965). The results from this estimation procedure are available in Table 7. We estimate mean fixed costs of about US\$102 and wholesale costs of approximately 60% of the manufacturer's suggested retail price.

[14]See Appendix E.1 for technical details on the BLP and MS fixed point iterations.

[15]DFO-LS is a derivative-free Gauss-Newton optimization method. It interpolates points to find an approximate Jacobian, constructs a locally quadratic model, and alternates minimizing the objective and updating the interpolation set.

| Parameter | Value | Description |
|---|---|---|
| $\theta_0^F$ | 9.23 [9.22,9.24] | mean of log fix cost |
| $\theta_M^F$ | 0.71 [0.69,0.72] | .. interacted with log market size |
| $\theta_R^F$ | 1.11 [1.10,1.13] | .. interacted with MSRP |
| $\theta_\sigma^F$ | 0.09 [0.08,0.10] | variance of log fix cost |
| $\theta_0^c$ | 90.17 [51.77,128.57] | coef. of marginal cost on const |
| $\theta_R^c$ | 0.60 [0.59,0.61] | coef. of marginal cost on MSRP |
| $\theta_\sigma^c$ | 0.01 [0.00,0.02] | variance of marginal cost |

**Table 7:** *SMM Parameters.*

Notes*: This table displays wholesale and fixed cost parameter estimates obtained from our simulated method of moments procedure. 95% confidence intervals are displayed under each parameter estimate.*

# 5 Counterfactuals

With our estimates in hand, we investigate two questions. Firstly, what is the overall effect of self-preferencing? To evaluate this effect, we compare factual outcomes to a counterfactual in which we set $\beta^r_{Amazon} = 0 < \hat{\beta}^r_{Amazon}$. Secondly, what is the combined effect of the platform's search guidance? To answer this question, we compare realized outcomes to a counterfactual in which naive consumers have to, absent recommendations, consider a random offer.

We distinguish three time-horizons in our counterfactuals. In the short run, we disallow sellers from changing both their entry decision and their prices. In the medium run, sellers price optimally but may neither enter nor exit their market. Finally, in the long run, sellers may change both their prices and entry decisions.

All results below refer to an estimation sample of 11,148 products, representing approximately US$24 million in revenue and 2.2 million customer arrivals per month. We know that "more than 1.9 million small and medium-sized businesses" list their products on Amazon, making up "close to 60% of [its] retail sales" (Bezos 2021). Furthermore, based on public disclosures, market research firms have estimated that the total revenue attributable to third-party sellers on Amazon Marketplace is $300 billion (Marketplace Pulse 2021). Thus, we can extrapolate our results to the entirety of Amazon Marketplace by multiplying our numbers by $300B/$24M $\approx$ 12 500. The resulting numbers give a good sense of the magnitude of the issues at stake. Thus, we use them to discuss our results, with the expected caveat that these naive extrapolations should be taken with a grain of salt.

## 5.1 "Self-dealing" can raise consumer welfare

We begin by investigating the overall effects of the recommendation algorithm's preference for the platform's own offers. To this end, we compare factual outcomes (where $\beta^r_{Amazon} = \hat{\beta}^r_{Amazon}$) to a counterfactual in which we set $\beta^r_{Amazon} = 0$. We report our results in Table 8, which summarizes our estimates of the causal effect of self-preferencing on the subset of the market which we observe. Thus, when the table reports, e.g., a \$312,047 figure for "Δ Total Consumer Surplus" in the "Short-Run" column, this indicates that total consumer surplus across the 11,148 products increases by this amount due to self-preferencing, holding pricing and entry fixed.

Starting our discussion with these short-run results, we find that the platform's recommendation advantage raises consumer welfare in the short run. We obtain this perhaps counterintuitive result because consumers prefer the platform's offers to those of third-party merchants, maybe because they are worried about counterfeit goods.

In the medium run, prices rise very slightly as self-preferencing de-emphasizes price as a recommendation selection criterion. This increase in prices harms all consumers but is offset by the static welfare gain for naive consumers. Sophisticated consumers, on the other hand, always ignore the recommendation and hence are only affected by its impact on prices.

Finally, we find a minimal effect of self-preferencing on entry by third-party merchants in the long run. While there is entry displacement, at 0.01 entrants/market, this effect is barely significant enough to register. This is because the platform's offers are already very attractive to consumers. Consumers prefer Amazon's offers in general and are also attracted to them because of their competitive prices. These effects ensure that the average merchant is already well aware that she will only be exposed to a small share of demand if she competes with the platform's own offers. Hence, the slight change in demand she is exposed to due to self-preferencing does not tilt the scales. Accordingly, the results of our counterfactual do not support the most plausible theory of harm from self-preferencing: decreased entry. Instead, extrapolated as discussed above, the overall consumer welfare gain from self-preferencing is \$4.5 billion.

However, we make two cautioning observations. Firstly, these results rely on our estimates of the extent to which consumers prefer Amazon offers to those of third-party merchants. As we do not have access to sales data for Amazon offers, we must infer these sales from estimates of market size and sales on other, non-Amazon offers. Although this yields no formal identification issues, there is room for future work to improve on our estimates with better consumer sales data.

Secondly, our model does not account for the platform's pricing strategy. In other words, we assume that Amazon's prices remain fixed throughout all counterfactuals. This assumption is necessary as we do not believe that modeling Amazon as maximizing short-run profits from item sales would be accurate. To compensate, we investigate in Appendix F.1 to what extent the platform would have to raise prices, as a consequence of its self-preferencing, to reverse our conclusion that this practice is welfare enhancing. This reversal happens if Amazon raises prices by more than 7.8%.

| Δ Outcome/Counterfactual | Short-Run | Medium-Run | Long-Run |
|---|---|---|---|
| Δ Platform Fees | -$47,669 | -$42,009 | -$41,295 |
| Δ Total Consumer Surplus | $312,407 | $386,787 | $398,131 |
| Δ Consumer Surplus (Naive) | $312,407 | $465,147 | $452,484 |
| Δ Consumer Surplus (Soph.) | $0 | -$78,359 | -$54,353 |
| Δ Producer Surplus | -$45,292 | $17,271 | $1,896 |
| Δ Welfare | $267,115 | $404,059 | $400,028 |
| Δ Mean (Price/MSRP) | 0.00% | 0.50% | 0.53% |
| Δ Mean # Entrants | 0.00 | 0.00 | -0.01 |

**Table 8:** *A Preference for the Platform's Own Offers Slightly Raises Welfare.*
Notes: *This table provides the difference in various outcomes that can be attributed to turning on the platform advantage. Recall $\beta^r_{Amazon}$ is the recommendation algorithm's coefficient on the dummy that indicates an offer belonging to the platform operator. Formally speaking, for various outcomes $x$, we compute $\Delta_x = \hat{x}(\hat{\beta}^r_{Amazon}) - \hat{x}(0)$ where $\hat{x}(\cdot)$ indicates our model prediction for an outcome as a function of $\beta^r_{Amazon}$. In the short run, sellers cannot change their prices or their entry decisions. In the medium run, sellers may set prices optimally, are not permitted to change their entry decisions. Finally, the long-run counterfactual allows sellers to change both their prices and their entry decisions.*

### 5.2 Search guidance benefits consumers

We now consider the combined effect of the search guidance algorithm. To this end, we compare market outcomes achieved by the estimated recommendation algorithm to those obtained when naive consumers consider offers uniformly at random. As these consumers can only consider one offer, this is a natural baseline for the case in which they are not provided with a recommendation. However, we emphasize that the counterfactual implicitly assumes that consumers would not adjust the size of their consideration sets in response to the absence of recommendations.[16]

We exhibit our results in Table 9, reporting the implied impact of the estimated search guidance algorithm on various market outcomes relative to the counterfactual described above. Beginning our discussion with the short-run results, we note that search guidance generates a large (extrapolated) static welfare gain of $50.9 billion. Furthermore, the platform achieves its triad of preferred outcomes: consumers benefit, merchants sell more, and the platform collects more fees. The intuition underlying this result is that offers on online marketplaces can vary dramatically in their attractiveness. Hence, it is essential to ensure that the 26% of consumers who can only consider one offer are successfully guided towards attractive offers.

However, the recommendation algorithm only achieves this matching of consumers to attractive offers by being extremely price-elastic. While this elasticity allows it to guide consumers towards competitively priced offers, merchants are also incentivized to attempt to capture a recommendation by underbidding each other. These incentives cause a 14.2% price decrease in the medium run, essentially redistributing surplus from the producer to the consumer side of the

[16]If this seems implausible, we note that the platform could always make its recommendations compulsory, i.e., only allow consumers to interact with recommended offers. Such a policy is implemented by other marketplaces such as ride-hailing services. In this context, our counterfactual world could be understood as one with a random recommendation that is compulsory only for naive consumers.

| Δ Outcome/Counterfactual | Short-Run | Medium-Run | Long-Run |
|---|---|---|---|
| Δ Platform Fees | $280,924 | -$550,486 | -$850,830 |
| Δ Total Consumer Surplus | $4,078,340 | $4,857,321 | $72,390 |
| Δ Consumer Surplus (Naive) | $4,078,340 | $1,148,209 | $967,462 |
| Δ Consumer Surplus (Soph.) | $0 | $3,709,112 | -$895,072 |
| Δ Producer Surplus | $152,173 | -$4,486,021 | -$1,193,062 |
| Δ Welfare | $4,230,513 | $371,299 | -$1,120,671 |
| Δ Mean (Price/MSRP) | 0.00% | -14.21% | -13.26% |
| Δ Mean # Entrants | 0.00 | 0.00 | -2.12 |

**Table 9:** *The Value of the Recommender System is Limited by its Long-Run Entry Implications.*
Notes*: This table provides the difference in various outcomes that can be attributed to the estimated recommendation algorithm's performance relative to a random baseline. Recall that $\boldsymbol{\beta^r}$ is the vector of weights that the recommendation algorithm places on various offer features when making its recommendation decision. Formally speaking, for various outcomes x, we report $\Delta_x = \hat{x}(\hat{\boldsymbol{\beta}}^{\boldsymbol{r}}) - \hat{x}(\mathbf{0})$. Setting $\boldsymbol{\beta^r} = \mathbf{0}$ is equivalent to the platform choosing recommendations uniformly at random. In the short run, sellers cannot change their prices or their entry decisions. In the medium run, sellers may set prices optimally, are not permitted to change their entry decisions. Finally, the long-run counterfactual allows sellers to change both their prices and their entry decisions.*

economy. At this point, consumers enjoy welfare gains of $60.7 billion from search guidance.

As might be expected, redistributing surplus in a two-sided market will have effects on entry. Indeed, we find that the platform's search guidance is responsible for a reduction of 2.21 entrants/market in the long run. Given that there are an average of 6 merchants on each market, this reduction indicates a significant dampening effect on entry. Our results suggest that this loss in entry results in major harm to the welfare of sophisticated consumers, who lose the ability to choose between various offers for the same product. The loss is sizeable because we estimate a prominent role for non-price determinants of demand and, in particular, for the unobserved offer quality. As entrants do not know their quality before entry, the entrants lost to the platform after introducing the recommendation algorithm have random draws from the quality distribution. Thus, the quality of the highest-quality choice available to consumers is lowered substantially.

It follows that, relative to what an A/B test would show, the value of the recommendation algorithm is much lower when entry is taken into account. Indeed, the entry effect lowers the (extrapolated) welfare gains from search guidance technology to $0.9 billion, a far cry from the medium-run gains of $60.7 billion. The effect of price-based search guidance on entry might thus explain why the platform has not chosen to make its recommendation algorithm even more price elastic. Indeed, in Appendix F.2, we find that compared to a world with more elastic recommendations, consumer welfare is higher under the currently employed algorithm.

## 6 Conclusion

Regulators have worried that platforms matching buyers and sellers may influence market outcomes by guiding consumer search through algorithmic recommendations. We have explored this issue by building a model of intermediation power. Our model is powerful enough to enable algorithmic influence on consideration sets, while remaining flexible enough to let the data speak to the extent of this influence. Our findings demonstrate the power of the default option. On Amazon marketplace, we find that 26% of customers only consider the default option, and the recommendation algorithm raises the price elasticity of demand from 3 to 4. It follows that through their choice of search, ranking, and recommendation algorithms, platforms are indeed able to influence market outcomes, guide consumers towards their own products and choose the intensity of price competition. In short, the power of defaults gives platforms a lot of intermediation power.

However, the sheer existence of intermediation power does not necessarily imply that platforms exploit this power in a way that is harmful to consumers. While we find that the examined recommendation algorithm advantages the platform, this self-preferencing has positive consequences for consumer welfare as consumers also prefer Amazon offers. Crucially, our model allows us to assess a theory of harm that has recently gained increasing attention, i.e. the idea that self-preferencing harms consumers by acting as an effective barrier to entrance. Our results do not support this conclusion.

Nevertheless, entry matters. Compared to a random recommendation baseline, the recommendation algorithm we examine brings large consumer surplus in the short and medium run. This surplus derives from the algorithms' ability to show consumers offers they are likely to purchase and its positive effect on the intensity of price competition. In the long run, however, this increase in competitive pressure lowers entry incentives. The reduction in entry, in turn, offsets nearly all of the algorithm's gains in consumer welfare relative to the random recommendation baseline.

One drawback of our current model is that we do not allow entry to be selective on merchant quality. While this extension is work in progress, we find that entry remains lowered with the recommender system relative to the random offer benchmark. Furthermore, future work on this issue could exploit our computationally tractable model while employing richer demand data if it becomes available. Such work could help assess the extent to which our attempt to compensate for the short-comings of our data by, e.g., proxying for market size has affected our conclusions. Finally, while platforms are large, they still have competitors. While our entry margin implicitly accounts for the ability of merchants to switch platform, a better though less tractable model would explicitly account for this dimension and allow both consumers and merchants to multi-home.

## References


Ackerberg, Daniel A (2009). "A new use of importance sampling to reduce computational burden in simulation estimation". In: *QME* 7.4, pp. 343–376.

Aguiar, Luis and Joel Waldfogel (2018). "Quality predictability and the welfare benefits from new products: Evidence from the digitization of recorded music". In: *Journal of Political Economy* 126.2, pp. 492–524.

Anderson, Simon P, Andre De Palma, and Jacques-Francois Thisse (1992). *Discrete choice theory of product differentiation*. MIT press.

Arellano, Manuel and Bo Honoré (2001). "Panel data models: some recent developments". In: *Handbook of Econometrics*. Ed. by J.J. Heckman and E.E. Leamer. 1st ed. Vol. 5. Elsevier. Chap. 53, pp. 3229–3296.

Bailey, Joseph P (1998). "Intermediation and electronic markets: Aggregation and pricing in Internet commerce". PhD thesis. Massachusetts Institute of Technology.

Baye, Michael R, John Morgan, and Patrick Scholten (2004). "Price dispersion in the small and in the large: Evidence from an internet price comparison site". In: *The Journal of Industrial Economics* 52.4, pp. 463–496.

Berry, Steven T (1992). "Estimation of a Model of Entry in the Airline Industry". In: *Econometrica*, pp. 889–917.

— (1994). "Estimating discrete-choice models of product differentiation". In: *The RAND Journal of Economics*, pp. 242–262.

Berry, Steven T and Peter C Reiss (2007). "Empirical models of entry and market structure". In: *Handbook of industrial organization* 3, pp. 1845–1886.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995). "Automobile prices in market equilibrium". In: *Econometrica*, pp. 841–890.

Berthene, April (2017). "Amazon's Prime Day conversion rate, traffic and transactions surge". In: URL: https://www.digitalcommerce360.com/2017/07/14/amazons-prime-day-conversion-rate-traffic-transactions-surge/.

Bezos, Jeff (2021). "2020 Letter to Shareholders". In: URL: https://www.aboutamazon.com/news/company-news/2020-letter-to-shareholders.

Bresnahan, Timothy F and Peter C Reiss (1991). "Entry and competition in concentrated markets". In: *Journal of Political Economy* 99.5, pp. 977–1009.

Brynjolfsson, Erik, Yu Jeffrey Hu, and Michael D Smith (2006). "From niches to riches: Anatomy of the long tail". In: *Sloan Management Review* 47.4, pp. 67–71.

Brynjolfsson, Erik, Yu Hu, and Michael D Smith (2003). "Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers". In: *Management Science* 49.11, pp. 1580–1596.

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

Cartis, Coralia et al. (2019). "Improving the flexibility and robustness of model-based derivative-free optimization solvers". In: *ACM Transactions on Mathematical Software (TOMS)* 45.3, pp. 1–41.

Casner, Ben (2020). "Seller curation in platforms". In: *International Journal of Industrial Organization* 72, p. 102659.

Caudet, Lucia and Maria Tsoni (2019). *Antitrust: Commission opens investigation into possible anti-competitive conduct of Amazon*. URL: https://europa.eu/rapid/press-release_IP-19-4291_en.htm.

Chamberlain, Gary (1980). "Analysis of Covariance with Qualitative Data". In: *Review of Economic Studies* 47.1, pp. 225–238.

Che, Yeon-Koo and Johannes Hörner (2017). "Recommender systems as mechanisms for social learning". In: *The Quarterly Journal of Economics* 133.2, pp. 871–925.

Chernozhukov, Victor, Han Hong, and Elie Tamer (2007). "Estimation and confidence regions for parameter sets in econometric models 1". In: *Econometrica* 75.5, pp. 1243–1284.

Chevalier, Judith A and Dina Mayzlin (2006). "The Effect of Word of Mouth on Sales: Online Book Reviews". In: *Journal of Marketing Research* 43.3, pp. 345–354.

Chunduru, Aditya (2020). *EU to force online platforms to make search ranking algorithms more transparent*. URL: https://www.medianama.com/2020/12/223-european-union-ranking-algorithms-transparency-online-platforms/.

Ciliberto, Federico and Elie Tamer (2004). "Market structure and multiple equilibria in airline markets". In: *Econometrica*.

Conlon, Christopher and Jeff Gortmaker (2020). "Best practices for differentiated products demand estimation with PyBLP". In: *RAND Journal of Economics* 51.4, pp. 1108–1161.

De los Santos, Babur, Daniel P O'Brien, and Matthijs R Wildenbeest (2018a). "Agency Pricing and Bargaining: Evidence from the E-Book Market". In: *Kelley School of Business Research Paper* 18-90.

De los Santos, Babur, Daniel P O'Brien, and Matthijs R. Wildenbeest (2018b). "Agency Pricing and Bargaining: Evidence from the E-Book Market". In: *papers.ssrn.com*.

Dinerstein, Michael et al. (2018). "Consumer price search and platform design in internet commerce". In: *American Economic Review* 108.7, pp. 1820–59.

District of Columbia (2021). *District of Columbia vs Amazon Inc*. URL: https://oag.dc.gov/sites/default/files/2021-05/Amazon-Complaint-.pdf.

Donnelly, Robert, Ayush Kanodia, and Ilya Morozov (2021). "The Long Tail Effect of Personalized Rankings". In: pp. 1–56.

Einav, Liran et al. (2015). "Assessing sale strategies in online markets using matched listings". In: *American Economic Journal: Microeconomics* 7.2, pp. 215–47.

Ellison, Glenn and Sara Fisher Ellison (2009). "Search, obfuscation, and price elasticities on the internet". In: *Econometrica* 77.2, pp. 427–452.

Gandhi, Amit, Zhentong Lu, and Xiaoxia Shi (2017). "Estimating Demand for Differentiated Products with Zeroes in Market Share Data". In: pp. 1–66.

German Competition Authority, Bundeskartellamt (2013). *Decision B 6 – 46/12*.

— (2015). *Decision B 9 – 121/13*.

Goeree, Michelle Sovinsky (2008). "Limited information and advertising in the US personal computer industry". In: *Econometrica* 76.5, pp. 1017–1074.

Gonzalez, Guadalupe (2018). "Phantom Buy Buttons, Performance Dings: The Scariest Things That Amazon Sellers Say Keep Them Up at Night". In: *Inc.* URL: https://www.inc.com/guadalupe-gonzalez/what-keeps-amazon-sellers-up-at-night-third-quarter-results.html (visited on 06/03/2021).

Hendricks, Kenneth and Robert H Porter (2007). "An empirical perspective on auctions". In: *Handbook of Industrial Organization* 3, pp. 2073–2143.

Janssen, Rebecca et al. (2021). *GDPR and the Lost Generation of Innovative Apps*. Working Paper.

Jovanovic, Boyan (1982). "Selection and the Evolution of Industry". In: *Econometrica: Journal of the Econometric Society*, pp. 649–670.

Kelly, Makena (2019). *Amazon silently ends controversial pricing agreements with sellers*. URL: https://www.theverge.com/2019/3/11/18260700/amazon-anti-competitive-pricing-agreements-3rd-party-sellers-end.

Klingler, Désirée et al. (2020). *Amazon's Theory of Harm*. Tech. rep. Yale University.

Levin, Dan and James L Smith (1994). "Equilibrium in auctions with entry". In: *The American Economic Review*, pp. 585–599.

Malmendier, Ulrike and Young Han Lee (2011). "The bidder's curse". In: *American Economic Review* 101.2, pp. 749–87.

Maplesden, Paul (2019). *How to Win The Amazon Buy Box: Your Questions Answered*. URL: https://www.webretailer.com/lean-commerce/amazon-buy-box/.

Marketplace Pulse (2021). *Marketplaces Year in Review*. Tech. rep. URL: https://www.marketplacepulse.com/marketplaces-year-in-review-2020.

Mattioli, Dana (2019). *Amazon Changed Search Algorithm in Ways That Boost Its Own Products*. URL: https://www.wsj.com/articles/amazon-changed-search-algorithm-in-ways-that-boost-its-own-products-11568645345?mod=rsswn.

Matzkin, Rosa L (2007). "Chapter 73 Nonparametric identification". In: Elsevier, pp. 5307–5368.

Mazzeo, Michael J (2002). "Product choice and oligopoly market structure". In: *RAND Journal of Economics*, pp. 221–242.

McFadden, Daniel (1981). "Econometric models of probabilistic choice". In: *Structural analysis of discrete data with econometric applications* 198272.

— (1989). "A method of simulated moments for estimation of discrete response models without numerical integration". In: *Econometrica: Journal of the Econometric Society*, pp. 995–1026.

Morrow, W Ross and Steven J Skerlos (2010). "Fixed-Point Approaches to Computing Bertrand-Nash Equilibrium Prices Under Mixed Logit Demand: A Technical Framework for Analysis and Efficient Computational Methods". In: *arXiv preprint arXiv:1012.5836*.

— (2011). "Fixed-point approaches to computing Bertrand-Nash equilibrium prices under mixed-logit demand". In: *Operations Research* 59.2, pp. 328–345.

Neidig, Harper (2019). "Tech giants on defensive at antitrust hearing". In: *The Hill*. URL: https://thehill.com/policy/technology/453390-tech-giants-on-defensive-at-antitrust-hearing.

Nelder, John A and Roger Mead (1965). "A simplex method for function minimization". In: *The Computer Journal* 7.4, pp. 308–313.

Nevo, Aviv (2001). "Measuring Market Power in the Ready-to-Eat Cereal Industry". In: *Econometrica* 69.2, pp. 307–342.

Pakes, Ariel and David Pollard (1989). "Simulation and the asymptotics of optimization estimators". In: *Econometrica* 57.5, pp. 1027–1057.

Piontek, Rachel (2019). "Committee for the Study of Digital Platforms Market Structure and Antitrust Subcommittee". In: pp. 1–100.

Quan, Thomas W and Kevin R Williams (2018). "Product variety, across-market demand heterogeneity, and the value of online retail". In: *The RAND Journal of Economics* 49.4, pp. 877–913.

Rasch, Georg (1960). *Studies in Mathematical Psychology: I. Probabilistic Models for Some Intelligence and Attainment Tests.* Nielsen & Lydiche.

— (1961). "On General Laws and the Meaning of Measurement in Psychology". In: *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*. Vol. 4, pp. 321–333.

Reimers, I and J Waldfogel (2019). "Digitization and Product Discovery: The Causal and Welfare Impacts of Reviews and Crowd Ratings". In: *American Economic Review*.

Roberts, James and Andrew Sweeting (2010). "Entry and Selection in Auctions". In:

— (2013). "When should sellers use auctions?" In: *American Economic Review* 103.5, pp. 1830–61.

Samuelson, William F (1985). "Competitive bidding with entry costs". In: *Economics Letters* 17.1-2, pp. 53–57.

Segarra, Lisa Marie (2018). *Google to pay Apple $12 billion to remain Safari's default search engine in 2019: Report*. URL: https://finance.yahoo.com/news/google-pay-apple-12-billion-155007747.html.

Smith, Michael D and Erik Brynjolfsson (2001). "Consumer decision-making at an Internet shopbot: Brand still matters". In: *The Journal of Industrial Economics* 49.4, pp. 541–558.

Stigler, George J (1961). "The economics of information". In: *Journal of Political Economy* 69.3, pp. 213–225.

Tamer, Elie (2003). "Incomplete simultaneous discrete response model with multiple equilibria". In: *The Review of Economic Studies* 70.1, pp. 147–165.

Thaler, Richard H. and Cass R. Sunstein (2008). *Nudge. Improving Decisions About Health, Wealth, and Happiness*. New Haven, CT and London: Yale University Press. 304 pp. ISBN: 978-0-300-12223-7.

UK Competition Authority, Office of Fair Trading (2013). *case CE/9692/12*.

U.S. Congress (2021). *American Choice and Innovation Online Act*. URL: https://www.congress.gov/bill/117th-congress/house-bill/3816/text.

U.S. Senate Judiciary Committee (2020). *Investigation of Competition in Digital Markets*. Tech. rep. Congress. URL: https://judiciary.house.gov/uploadedfiles/competition_in_digital_markets.pdf.

## Part

# Appendix

## Table of Contents

# A Data Description

## A.1 Recommendation Status Data

Our primary data source is a repricing company that administrates offer listings on behalf of third-party merchants on Amazon Marketplace. The company provides us with three data sets of interest, both spanning one and a half years. Firstly, we have access to the notifications sent to the company programmatically by Amazon. These notifications are sent if there are any changes on any of the offers on the listings that the company administrates (e.g., updates its price, exits or enters). Secondly, the company is informed of any sales that occur on its customers' listings. Finally, we have regular data on the sales ranks of each product (i.e., market, not offer); we discuss this last dataset in more detail in Appendix D.1 below.

To begin with, the notifications are our primary source of information about the alternatives available on any given market at any given time. For each notification and each offer, we observe (i) the recommendation status, (ii) whether the offer belongs to a "featured" merchant, (iii) whether Amazon fulfills it, (iv) its cost and (v) shipping price, (vi) the feedback count and (v) fraction of positive feedback for the associated seller, and (vi) the time to dispatch. The main challenge with using this dataset is the uneven time resolution: a notification is sent whenever any offer characteristic *other than recommendation status* changes. Thus, while we have infinite temporal resolution on, e.g., prices, we lack this precision in recommendation status.

The demand data contain, for each merchant registered with the repricing company, the exact time of every sale the merchant made during its period of registration. The main shortcoming of this dataset is that we cannot independently verify when a merchant is registered with the company.

We combine these datasets using the following procedure. First, we filter the notification data to list only merchants that are "featured." This is a hard prerequisite for a merchant to become eligible to be recommended. To become a "featured" merchant (FM), the seller has to pay $40 a month to Amazon for a professional seller account. Second, we only retain products in the "Fashion" category. Third, for each product, we proxy the period that the seller of said product was registered with the repricing company. To do so, we find the dates of the first and last observed sale; then, we declare the product *observed* on any dates between those two. Finally, we merge our data by aggregating up both the notification and sales information to the product-day level by taking averages. Recall that, whenever any offer characteristics change, we receive a notification. Therefore, on each product, we forward fill from the last observed notification for any days without any notifications. After we have merged our data, we obtain one dataset at the offer-date level with 204,017,670 observations. Summary statistics for this dataset are provided in Table 12.

We summarize the data using a correlation heatmap in Figure 6 below. Offers shipped by Amazon ship faster (as they are always dispatched instantly) and do not charge for shipping. Furthermore, sellers who employ FBA typically receive more feedback. More importantly, the


**Figure 6:** *Correlation Heat Map for AOCs.*
Notes*: This figure illustrates the correlations between various offer-level covariates. A larger square indicates a larger correlation, and the color of the square indicates both size of the correlation as well as direction (red for negative, green for positive).*

raw correlations reveal that on average, Amazon is more likely to win the Buybox whenever it competes. Furthermore, Buybox status depends on shipping time and — as expected — listing and shipping price.

As the consumers always pay the "landed" price, i.e., listing price plus shipping cost, we combine these variables and refer to their sum as the *price* from now on. Next, we investigate how, within a market, recommendation status depends on the rank of an offer. To this purpose, Figure 3a illustrates that there is a considerable amount of price variation across offers for the same product (left panel) and that landed price is one of the primary determinants for recommendation status (right panel).

## A.2 Sales Data

After subsetting to the products in the 'Fashion' category, we observe a total of 256,704 sales across 11,148 products over the period from August 25, 2018 to March 26, 2020. The number of observed sales varies widely product-by-product as illustrated in Figure 8a. The minimum number of sales was attained by a particular women's boot, which only sold twice during our

| | Avg. | Std. Dev. | Min. | 25th | 50th | 75th | Max. |
|---|---|---|---|---|---|---|---|
| ListingPrice | 58.08 | 44.50 | 0.00 | 24.08 | 44.96 | 82.32 | 9999.00 |
| % Positive Feedback | 96.04 | 10.00 | 0.00 | 96.00 | 99.00 | 100.00 | 100.00 |
| IsFulfilledByAmazon | 0.80 | 0.40 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Is Buybox Winner? | 0.11 | 0.31 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |
| Feedback Count | 8.38 | 38.12 | 0.00 | 0.25 | 0.81 | 2.88 | 3023.68 |
| Shipping Speed | 11.09 | 33.12 | 0.00 | 0.00 | 0.00 | 0.00 | 1620.00 |
| Shipping | 0.38 | 1.86 | 0.00 | 0.00 | 0.00 | 0.00 | 100.00 |
| IsFeaturedMerchant | 1.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| IsAmazon | 0.01 | 0.11 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |

**Table 10:** *Offer-Level Summary Statistics.*

Notes: *This table displays summary statistics for the dataset used in recommendation and consumer choice estimation. Note that "time to dispatch" is measured in hours, and a time to dispatch of zero means that the product is shipped as soon as it is ordered. Meanwhile, "Feedback Count" is measured in thousands.*

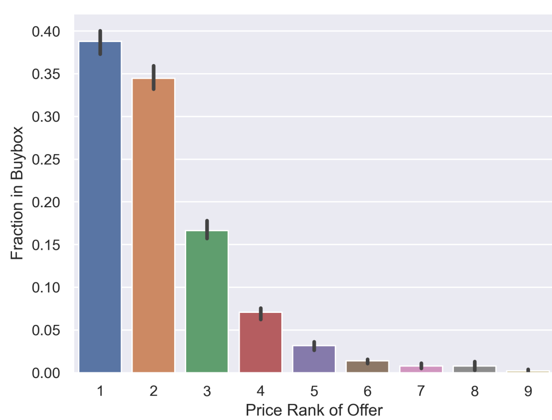


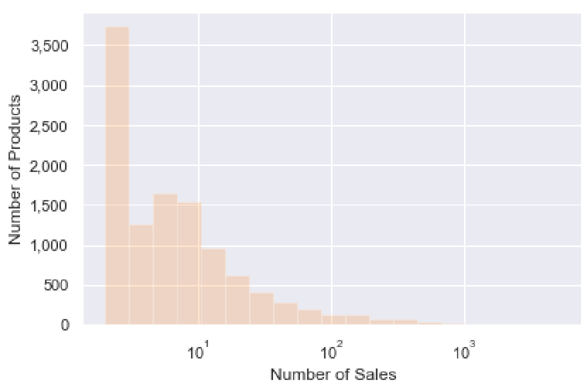

**Figure 7:** *Recommendation Status Varies With Price Rank (Amazon Markets).*

Notes: *This figure displays the fraction of offers at a given price rank that are recommended, restricting our attention to only the markets in which Amazon is competing. Black bars represent 95% confidence intervals from mean estimation.*

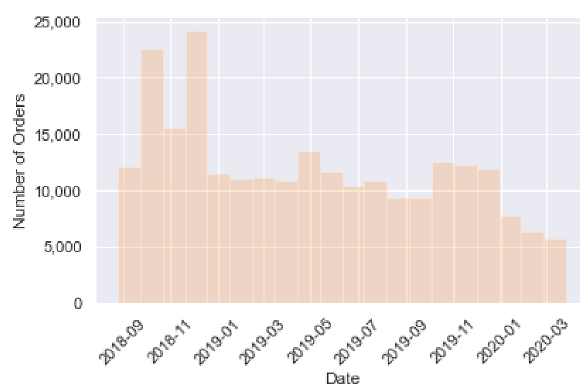
| Variable | Coefficient |
|---|---|
| $1\{\text{Product Price above MSRP}\}$ | -0.02 [-0.03, -0.01] |
| $\log\left(\frac{\text{Product Price}}{\text{MSRP}}\right)$ | 0.09 [0.07,0.11] |
| $1\{\text{Product Price above MSRP}\} \times \log\left(\frac{\text{Product Price}}{\text{MSRP}}\right)$ | 0.64 [0.57,0.71] |
| Product FE? | ✓ |

**Table 11:** *Parameter Estimates Corresponding to the Canceled Recommendation Graph.*
Notes: *This table provides the estimates of running model (1) in the main text. Here, we define the product price to be the minimum price across all offers. We find no evidence of a positive discontinuity around MSRP, but strong evidence that (i) a higher price increases the probability of a cancelled recommendation everywhere (as $\beta_1 > 0$ and $\beta_2 > 0$) and (ii) this effect is much stronger when the price exceeds MSRP.*

sampling timeframe. The maximum number of sales was obtained by "Party Sunglasses", selling 5,376 times. We observe significantly more sales before the start of 2019 as our data vendor lost access to data on a large number of products when one of its customers left Amazon in January 2019. Our total distribution of sales over time suggests no (further) anomalous events (see Figure 8b).




**(a)** *Distribution # Sales By Product.*




**(b)** *Distribution # Orders Over Time.*

**Figure 8:** *Distribution of Sales By Product and Time.*
Notes: *This figure illustrates the long-tail nature of our sales data (left panel) and confirms that we missed no anomalous events (right panel).*

### A.3 Keepa Data

We obtain further data from a broker that sells access to details about many Amazon Marketplace products, including, importantly for us, the manufacturer's suggested retail price (MSRP). In particular, we queried the Keepa[17] database for each product used in our demand estimation from May 12, 2020 to August 11, 2020. We use this dataset to (i) categorize each product into one of Amazon's top-level product categories and (ii) to associate each product with a unique MSRP.

This data set contains 18,873,203 observations over 1,663,789 "snapshots." Each snapshot is a product-time pair. An observation is a seller-snapshot pair. The observations are spread over 13,138 products and 448,923 unique time stamps.

| | Avg. | Std. Dev. | Min. | 25th | 50th | 75th | Max. |
|---|---|---|---|---|---|---|---|
| isAmazon | 0.02 | 0.13 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |
| isFBA | 0.65 | 0.48 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 |
| isPrime | 0.67 | 0.47 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 |
| p | 5734.57 | 4926.00 | 439.00 | 2357.00 | 3894.00 | 7735.00 | 34899.00 |
| r | 0.09 | 0.29 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |
| AsinId | 6922.97 | 4036.36 | 0.00 | 3271.00 | 6928.00 | 10413.00 | 13800.00 |
| MerchantId | 9827.60 | 5743.13 | 0.00 | 4721.00 | 9664.00 | 14999.00 | 19240.00 |

**Table 12:** *Offer-Level Summary Statistics for Keepa Data.*
Notes*: This table displays summary statistics for the auxiliary dataset used to recover Manufacturer Suggested Retail Prices. Note that "isFBA" is an indicator for whether the merchant uses Amazon's dispatch service.*

Similarly to our main repricer data set, we begin by investigating the correlations between these variables. The correlation heatmap (Figure 9) corroborates the results in the previous section. Amazon appears more likely to recommend its own offers, and merchants who fulfill their goods using Amazon's logistics service are more likely to be Prime merchants.

We now demonstrate that there is a considerable amount of price variation across offers. This variation is shown in Figure 10. Similarly to Figure 3, we find that prices vary substantially for the same product (left panel); furthermore, landed price remains one of the primary determinants for recommendation status (right panel). However, compared to the data from our repricer, there appears to be more markup dispersion in the Keepa data. Specifically, maximal markups can be 2.5 times relative to the lowest-priced offer (relative to 2 times in the main data set). On the other hand, the lowest-priced offer is only recommended less than 40% of the time in the Keepa data, whereas it is recommended over 55% of the time in our main data set.

Finally, we extend the analysis in Table 2 to our auxiliary data set. These results are displayed in Table 13. As in the mainline descriptive analysis, we note that only 37.08% of the lowest-priced offers in our sample are recommended by Amazon. However, the recommended offer is fulfilled by Amazon (FBA) in 76.58% of cases, a much higher number. Meanwhile, Amazon recommends

[17] http://keepa.com

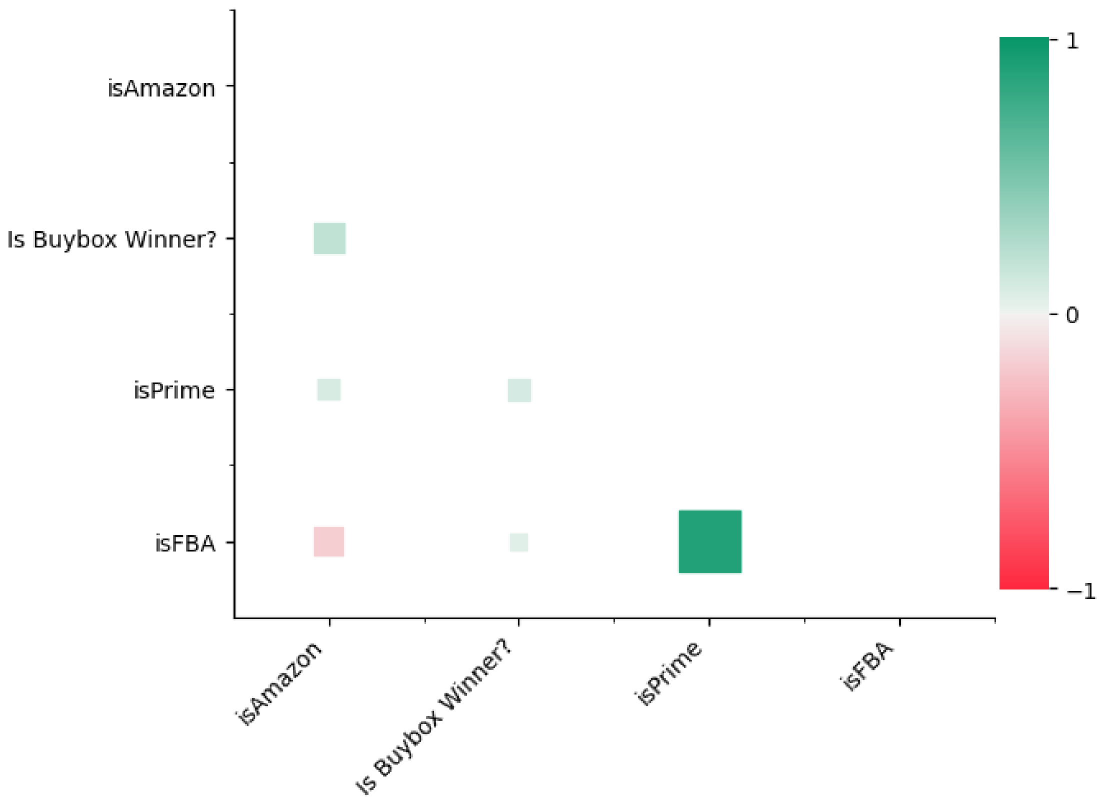
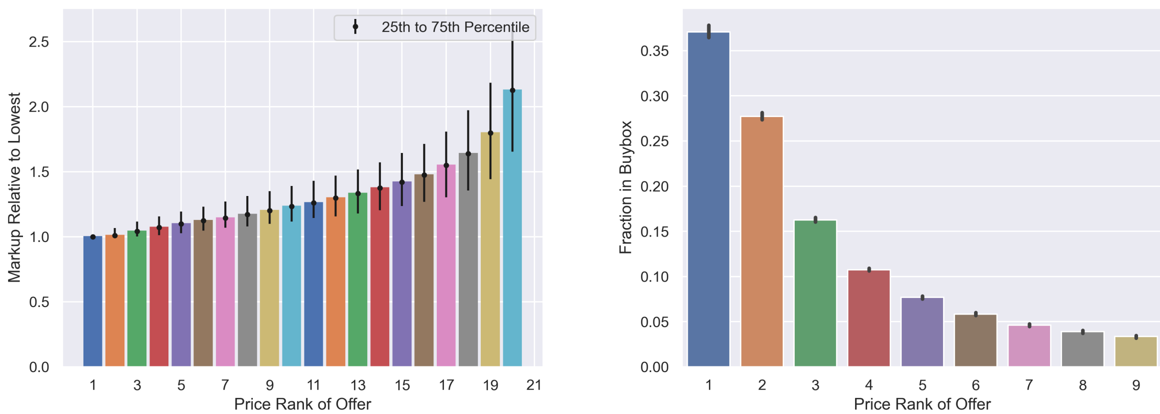



**Figure 9:** *Correlation Heat Map for AOCs in the Keepa data.*
Notes*: This figure illustrates the correlations between various offer-level covariates. A larger square indicates a larger correlation, and the color of the square indicates both size of the correlation as well as direction (red for negative, green for positive).*




**(a)** *Price Multiple.* **(b)** *Mean Recommendation Status.*

**Figure 10:** *Prices and Recommendation Status Vary Significantly with Price Rank (Keepa Data).*
Notes*: The left figure displays the median markup of the n-th lowest-priced offer over the lowest-priced offer on each product, for various price ranks n. The right figure gives the fraction of offers at a given price rank that are recommended. Black bars in the left figure indicate interquartile ranges, while those on the right are 95% confidence intervals from mean estimation.*

its own offers about 10.20% of the time, but the company does not list an offer on most of the products in this sample. When we condition on products with an Amazon offer (i.e. the rightmost column), the platform recommends its own offer 54.52% of the time. In conclusion, where we have data on our covariates in both our main and auxiliary data sets, the results from the latter corroborate those of our mainline analysis.

| | Fraction of Observations | | |
|---|---|---|---|
| | Overall | FBA | Amazon |
| Amazon Offer Exists | 18.71% | 17.43% | 100.00% |
| FBA Offer Exists | 92.97% | 100.00% | 86.63% |
| Winner Is ... | | | |
| Lowest Priced | 37.08% | 34.63% | 33.75% |
| Second Lowest Priced | 25.64% | 26.22% | 24.78% |
| Lowest Priced FBA | 34.49% | 37.07% | 17.28% |
| FBA | 76.58% | 82.38% | 37.58% |
| Amazon | 10.20% | 8.72% | 54.52% |

**Table 13:** *Determinants of Recommendation Status (Keepa Data).*
Notes*: This table provides the fraction of recommendation observations satisfying the criteria listed in the first column. The second column is based on the entire set of observations, the third column employs only observations for which there is an offer that is fulfilled by Amazon and the fourth column uses only observations for which there is an offer by Amazon itself.*

# B Suggestive Evidence of Randomization

We model the platform's algorithm as solving a discrete-choice problem with extreme-valued shocks. These shocks are ascribed to deliberate randomization or mistakes in evaluating a seller's offer. In this section, we provide some suggestive evidence of the former. Even if seller offer characteristics remain fixed, platforms may not always recommend the same product, e.g., the cheapest one. To motivate this modeling choice, we select a product — Maxell Lithium Batteries — sold on the Amazon marketplace. In Figure 11, we display the identity of the merchant whose offer was recommended by the platform on August 14-15, 2019. The distinct merchants are colored red, green, and blue, respectively; their prices are fixed at \$2.67, \$2.70, and \$2.72 throughout the displayed period. We find that the platform "cycles" between the three sellers, giving each of them one-third of the share of the platform's recommendation. This finding is corroborated by various sites advising sellers on pricing strategies. For instance, Maplesden (2019) notes that

> "Amazon does seem to try and rotate the Buy Box between different sellers [. . . ] but it will weight the rotation more heavily towards the 'better' sellers."

Therefore, we build platform randomization into our model.

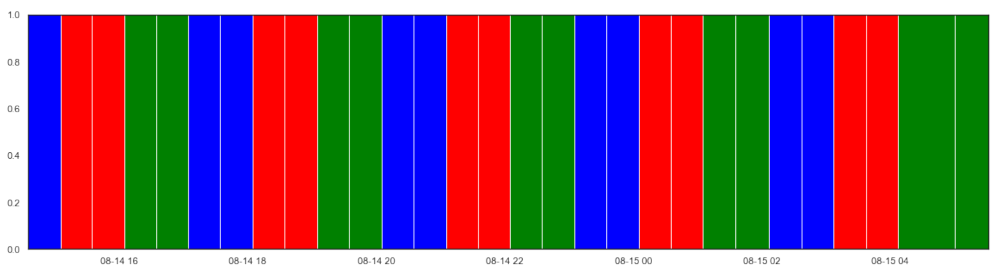



**Figure 11:** *The Buybox Rotation.*
Notes*: Time periods in which three distinct sellers of Maxell Lithium Batteries, colored red, green and blue, are recommended on the Amazon third-party platform. All merchants held their prices constant from August 14, 2019 at 2.30pm to August 15, 2019 at 5.30am. The prices set were \$2.67, \$2.70 and \$2.72 respectively.*

# C Model Details

## C.1 Entry Game Proofs

**Lemma 1.** *$V$ is continuous in the putative cutoff $c$.*

*Proof.* Write expected gross profits as

$$V(c) = \int \pi_j(c, (c_k)_{k \in K(c)}) f(c_{-j})\, dc_{-j}, \quad \text{where } K(c) = \{k \in \mathcal{N} \setminus \{j\} : c_k \leq c^*\}.$$

We note that $\pi_j$ is piecewise-continuous in each of $j$'s opponents' marginal costs, and the lognormal probability density function is continuous. Since the (definite) integral of a piecewise-continuous function is continuous, we have that $V$ is continuous in $c$. □

**Proposition 2.** *Assume that fixed costs are not too high, i.e. $V(c) \geq F$ for some $c \in [0, \infty)$. Then the entry game with heterogeneous marginal costs has a symmetric equilibrium in cutoff strategies.*

*Proof.* As $c^* \to \infty$, an arbitrarily large number of firms enter in equilibrium, driving profits down to zero, ceteris paribus. On the other hand, as $c^* \to 0$, if our assumed distribution for costs has no mass point at zero, only firm $j$ enters the market, setting the monopoly price and attaining monopoly profits. If fixed costs $F$ are less than monopoly profits, by the Intermediate Value Theorem, $\exists\, c^{**}$ for which equation $V(c^{**}) = F$ holds. □

**Lemma 2.** *Expected gross profits $V$ are monotonically decreasing in the marginal cost cutoff $c$.*

We relegate the proof of Lemma 2 to Subsection C.2 below. Here, we will give a heuristic argument in the case of two potential entrants. Fix a seller $j$ at the cost cutoff $c^*$. From her perspective, she faces competition when her opponent has a cost $c \leq c^*$, and she is a monopolist otherwise. Hence expected gross profits can be decomposed as

$$V(c^*) = \int_0^{c^*} \pi^T(c^*, c) f(c)\, dc + [1 - F(c^*)]\pi_M^T(c^*),$$

where $\pi^T$ represents a particular realization of gross profits conditional on entry:

$$\pi^T(c^*, c) = \pi_j(c^*, c)$$

and $\pi_M$ denotes the period profits of a monopolist with marginal cost $c^*$. Taking partial derivatives with respect to $c^*$, we get

$$\begin{aligned}
\frac{\partial V}{\partial c^*} &= \pi^T(c^*, c^*) f(c^*) + \int_0^{c^*} \left.\frac{\partial \pi^T(\tilde{c}, c)}{\partial \tilde{c}}\right|_{\tilde{c}=c^*} f(c)\, dc - f(c^*)\pi_M^T(c^*) + [1 - F(c^*)]\pi_M^{T\prime}(c^*). \\
&= f(c^*)[\pi^T(c^*, c^*) - \pi_M^T(c^*)] + \int_0^{c^*} \left.\frac{\partial \pi^T(\tilde{c}, c)}{\partial \tilde{c}}\right|_{\tilde{c}=c^*} f(c)\, dc + [1 - F(c^*)]\pi_M^{T\prime}(c^*).
\end{aligned}$$

We note that each of the three summands on the RHS are individually less than 0. This is because profits under competition are strictly less than under monopoly, and a seller's profits fall as her marginal costs rise. Hence $V$ is monotonically decreasing in $c$. Together with Proposition 2, this immediately implies

**Proposition 3.** *The entry game has a unique symmetric equilibrium in cutoff strategies.*

## C.2 Entry Game: Proof of Lemma 2

*Proof.* Fix a seller in a market with $N+1$ players (i.e. she has $N$ opponents). Her expected gross profits from entry at $c^*$ is

$$V(c^*) = [1-F(c^*)]^N \pi_M^T(c^*) + \sum_{n=1}^{N} \binom{N}{n} [1-F(c^*)]^{N-n} \int_{[0,c^*]^n} \pi_n^T(c^*, c_{-j}) f(c_{-j})\, dc_{-j},$$

where $\pi_n^T$ denotes her total profit function, after entry has resolved, with $n$ firms that have marginal cost draws less than $c^*$. Since her opponents draw their costs iid, $\pi_n^T$ is exchangeable in opponent costs for all $n$; hence we write

$$I_{k,n} := \int_{[0,c^*]^{n-k}} \pi_n^T(c^*, c_{1:n-k}, c_{n-k+1:n})\, dF(c_{1:n-k})$$

for the expectation of profits with respect to the "first $n-k$ opponents", taking the costs of the last $k$ opponents as fixed. Where required, we write

$$I_{k,n}(c_{n-k+1}, c_{n-k+2}, \ldots, c_n)$$

for the integral $I_{k,n}$ given that the last $k$ opponents have marginal costs $(c_{n-k+1}, c_{n-k+2}, \ldots, c_n)$. The partial derivative of expected gross profits can then be written as

$$\begin{aligned}\frac{\partial V}{\partial c^*} =& [1-F(c^*)]^N \pi_M^{T\,\prime}(c^*) - N[1-F(c^*)]^{N-1} f(c^*) \pi_M^T(c^*) + \\ & \sum_{n=1}^{N} \binom{N}{n} \Big\{ -(N-n)[1-F(c^*)]^{N-n-1} f(c^*) I_{0,n} \mathbf{1}\{n<N\} + \\ & [1-F(c^*)]^{N-n} \frac{\partial I_{0,n}}{\partial c^*} \Big\}. \end{aligned} \tag{C1}$$

To bound this partial derivative further, we prove three sublemmata:

1. If opponents are symmetric, the partial derivative of the integral $I_{0,n}$ with respect to $c^*$ is

$$\frac{\partial I_{0,n}}{\partial c^*} = \int_{[0,c^*]^n} \frac{\partial \pi_n^T(c^*, c_{-j})}{\partial c^*}\, dF(c_{-j}) + n f(c^*) I_{1,n}(c^*). \tag{C2}$$

   To see why this is the case, with some abuse of notation, let $F_{-m}$ be the probability measure under which players $j$ and $m$ are excluded. Apply Leibniz's rule to get

$$\frac{\partial I_{0,n}}{\partial c^*} = \int_0^{c^*} f(c_n) \frac{\partial}{\partial c^*} \int_{[0,c^*]^{n-1}} \pi_n^T\, dF_{-n}\, dc_n + f(c^*) I_{1,n}(c^*).$$

   Apply Leibniz's rule again to get

$$\frac{\partial I_{0,n}}{\partial c^*} = \int_0^{c^*} f(c_n)\Bigg[\int_0^{c^*} f(c_{n-1})\frac{\partial}{\partial c^*}\int_{[0,c^*]^{n-2}} \pi_n^T \, dF_{-n,-(n-1)}\, dc_{n-1}$$
$$+ f(c^*)\int_{[0,c^*]^{n-2}} \pi_n^T(c^*, c_1, \ldots, c_{n-2}, c^*, c_n)\, dF_{-n,-(n-1)}\Bigg] dc_n$$
$$+ f(c^*) I_{1,n}(c^*). \tag{C3}$$

Consider the term on the second line of Equation (C3). By exchangeability and Fubini's Theorem, we can rewrite it as

$$f(c^*)\int_0^{c^*} f(c_{n-1})\int_{[0,c^*]^{n-2}} \pi_n^T(c^*, c_1, \ldots, c_{n-2}, c_{n-1}, c^*)\, dF_{-n,-(n-1)}\, dc_{n-1},$$

which is precisely $f(c^*)I_{1,n}(c^*)$. Hence

$$\frac{\partial I_{0,n}}{\partial c^*} = \int_{[0,c^*]^2} \frac{\partial}{\partial c^*}\left[\int_{[0,c^*]^{n-2}} \pi_n^T \, dF_{-n,-(n-1)}\right] dF(c_{n-1}, c_n) + 2f(c^*)I_{1,n}(c^*).$$

Repeat this procedure another $n-2$ times. Since there are a total of $n$ integral signs with which to apply Leibniz's rule, we have that Equation (C2) holds, as desired.

2. $\binom{N}{n+1}(n+1) = \binom{N}{n}(N-n)$. This follows from

$$\frac{N!}{(n+1)!(N-n-1)!}(n+1) = \frac{N!}{n!(N-n-1)!} = \frac{N!}{n!(N-n)!}(N-n).$$

3. For any $n \in \mathbb{Z}_0^+$, we have

$$I_{0,n} > I_{1,n+1}(c^*). \tag{C4}$$

This follows as, fixing a realization of opponent costs, our seller's profits decline if there is an additional entrant with cost $c^*$, all else equal. Since this is true for all realizations of opponent costs in $[0, c^*]^n$, Inequality (C4) follows.

Now we return to Equation (C1). By the first sublemma, expand the partial derivatives $\frac{\partial I_{0,n}}{\partial c^*}$ by repeated applications of Leibniz's rule:

$$\begin{aligned}\frac{\partial V}{\partial c^*} =& [1-F(c^*)]^N \pi_M^{T\prime}(c^*) - N[1-F(c^*)]^{N-1}\pi_M^T(c^*)f(c^*) - \\ & \sum_{n=0}^{N-1}\binom{N}{n}(N-n)[1-F(c^*)]^{N-n-1} f(c^*) I_{0,n} + \\ & \sum_{n=1}^{N}\binom{N}{n}[1-F(c^*)]^{N-n}\left\{\int_{[0,c^*]^n}\frac{\partial \pi_n^T}{\partial c^*}\, dF(c_{-j}) + nf(c^*)I_{1,n}(c^*)\right\}.\end{aligned}$$

Since our seller's profits must decline in her costs, we can drop the first term (with $\pi_M^{T\prime}(c^*)$), as well as the innermost terms of the form

$$\int_{[0,c^*]^n} \frac{\partial \pi_n^T(c^*, c_{-j})}{\partial c^*} \, dF(c_{-j}).$$

Changing limits of summation, the partial derivative of $V$ is bounded above by

$$\sum_{n=0}^{N-1} [1 - F(c^*)]^{N-n-1} f(c^*) \left[ \binom{N}{n+1} (n+1) I_{1,n}(c^*) - \binom{N}{n} (N-n) I_{0,n} \right].$$

Now, by the second sublemma, we have that $\binom{N}{n+1}(n+1) = \binom{N}{n}(N-n)$. Hence the bound on the partial derivative of $V$ evaluates to

$$\sum_{n=0}^{N-1} [1 - F(c^*)]^{N-n-1} f(c^*) \binom{N}{n+1} (n+1) \left[ I_{1,n}(c^*) - I_{0,n} \right],$$

which we claim is negative. This follows because the coefficient of $I_{1,n}(c^*) - I_{0,n}$ is positive; moreover, by the third sublemma, $I_{1,n}(c^*) - I_{0,n}$ itself is negative. Hence $V$ is monotonically decreasing in $c^*$. □

# D Demand & Consumer Choice Estimation

## D.1 Market Size

Before we consider how the total demand for a product is split between its various offers, we need to ascertain its market size $M$. There are two issues in this estimation, one standard and one less so.

The more routine problem is translating overall sales for a product (once such a number has been obtained) into an assessment of the associated market size. Doing so requires taking a stance on the number of consumers who contemplate purchasing a given product but ultimately decide not to pull the trigger (i.e., choose the outside option). Conveniently, there is an intuitive way of performing this translation in online sales: the "conversion rate," i.e., the fraction of product site visits that turn into a sale. While we do not have data on the conversion rates specifically for the products in our sample, we calibrate the conversion rate to 12.3% as this is the average across products on Amazon (Berthene 2017).

The larger challenge is that for each product page $p$ in our dataset, we only observe sales of one alternative $j^o(p) \in \mathcal{J}_p$, the so-called "source" merchant.[18] As we are estimating an entry model, it is crucial for us to recover some measure of the scale of demand. If we only see the realized number of purchases from one merchant, we may obtain a poor proxy for the mean number of purchases.[19] To combat this problem, we exploit a finding in Chevalier and Mayzlin

[18] Because it is qua this merchant that we observe the data for this product, i.e., they are (indirectly) the data source.

[19] This is essentially the problem discussed in Gandhi, Lu, and Shi (2017), only that it arises for us mostly because of

(2006), built on in De los Santos, O'Brien, and Wildenbeest (2018b) and Reimers and Waldfogel (2019): (log) sales ranks are excellent proxies for (log) sales.

Denote by $A_{p,\tau}$ the number of sales for product $p$ in month[20] $\tau$. Instead of observing $A_{p\tau}$, we observe the sales for our observed merchant:

$$a_{p,j^o(p),\tau} = \frac{s_{p,j^o(p),\tau}}{1 - s_{p,0,\tau}} \times A_{p\tau}.$$

Here, $s_{pj\tau}$ refers to the market share of alternative $j$ on product $t$ and time $\tau$ (as in the main text). We will assume that[21]

$$\ln A_{p\tau} = \gamma_0 + \gamma_1 \ln(\text{rank}_{p\tau}).$$

This suggests estimating the following Poisson model of observed monthly sales against sales rank:[22]

$$\mathbb{E}[a_{p,j^o(p),\tau}|X] = \exp\Big(\gamma_0 + \gamma_1 \ln(\text{rank}_{p,\tau}) + \ln\big(\frac{s_{p,j^o(p),\tau}}{1 - s_{p,0,\tau}}\big)\Big).$$

However, this model is infeasible as we do not observe the fraction of sales $s_{p,j^o(p),\tau}$ that go to $j^o(p)$ (we only observe the number). Instead, we replace $\ln\big(\frac{s_{p,j^o(p),\tau}}{1-s_{p,0,\tau}}\big)$ with a linear function of the (empirical) fraction of time that the offer is recommended, $\theta_3 \log(r_{p,j^o(p),\tau})$, which is a plausible proxy for the former.

We further refine our estimation by restricting attention to product-months (i) for which our source has inventory on at least 80% of days, (ii) for which we have at least three sales rank observations, and (iii) for which our data source was in the Buybox at least 10% of the time. We also account for the presence of "variation" products (e.g., a shoe with multiple different sizes). For these products, the sales rank is based on the total number of sales across all variations, but our sales data is based on the specific variations sold by our seller. Finally, we add product-level FE to identify the slope of the 'log rank' vs. 'log sales' relationship purely off within-product time variation in sales.

Having estimated the feasible model, we can now form an estimate of the number of sales on a product:

$$\hat{A}_p = \exp\left[\hat{\gamma}_0 + \hat{\gamma}_1 \ln\left(\frac{1}{T_p}\sum_{\tau=1}^{T_p} \text{rank}_{p,\tau}\right)\right].$$

Here, we substitute in the average sales rank of a given product over time, then set $r = 1$ to predict sales for a merchant. Thus, our prediction is made as if said merchant were the single merchant on a product and hence recovers an estimate of total sales. The resulting numbers are less noisy estimates of $A_p$ than those we would have obtained had we not used sales rank data.

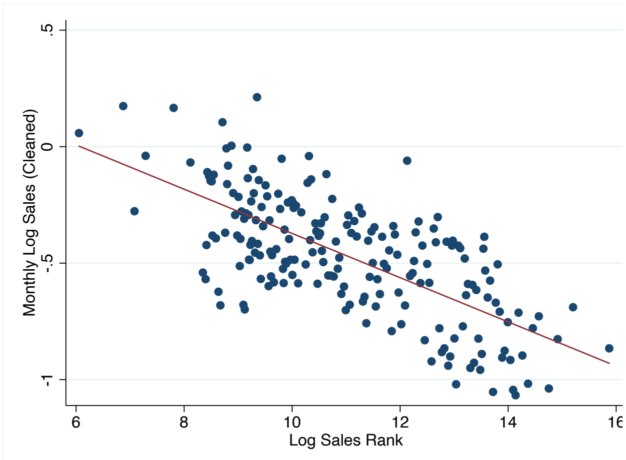
high-frequency data and partial observability, and that sales rank data allows us to employ a more tailor-made fix.

[20]Note that we aggregate to the monthly level for this estimation; in the main text, $\tau$ refers to a day.

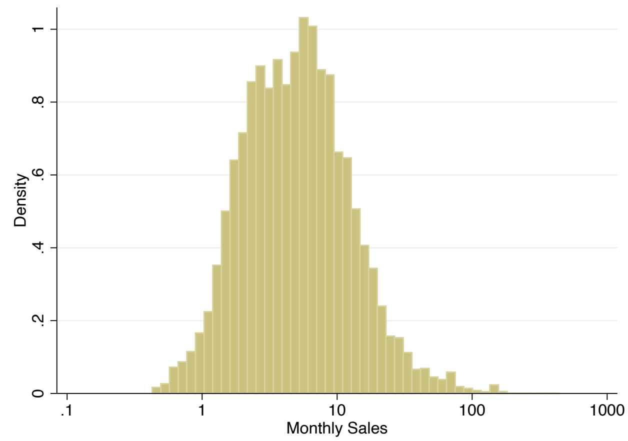
[21]Note that for ease of exposition, we pretend that there is only one category with respect to which all products are ranked here; in truth, we perform this estimation separately for each category.

[22]We prefer a Poisson model over the linear model in logs as it correctly deals with the frequent observations of zero sales on a given day.

We illustrate our procedure in Figure 12. The left panel shows a binned scatterplot of log sales rank cleaned from product-level fixed effects (on x) against monthly log sales (on y).[23] As expected, the linear model appears to fit the data well. The right panel shows predicted monthly sales. These sales vary a lot: while our sample includes products with an estimated 1000 sales per month, most products belong to the "long tail": they see few expected sales.




**(a)** *Relationship roughly linear.*




**(b)** *Plausible inferred sales.*

**Figure 12:** *Log Sales Ranks Predict Monthly Log Sales.*
Notes: *We use within-product variation in sales ranks and sales for products for which we observe the majority of demand to infer monthly sales for all products.*

Finally, we translate our estimate of monthly sales into an estimate of market size by calibrating the share of visits that result in a sale to 12.3%:

$$\hat{M}_p = \frac{1}{0.123}\hat{A}_p.$$

## D.2 MSL Estimation of Recommendation Algorithm

Recall from the main text that the probability that alternative $j$ on market $t$ is recommended is given by

$$r_{jt} = \frac{\left[\sum_{k\in\mathcal{J}_t}\exp(\delta^r_{kt}/\lambda)\right]^\lambda}{1+\left[\sum_{k\in\mathcal{J}_t}\exp(\delta^r_{kt}/\lambda)\right]^\lambda} \times \frac{\exp(\delta^r_{jt}/\lambda)}{\sum_{k\in\mathcal{J}_t}\exp(\delta^r_{kt}/\lambda)},$$

where

$$\delta^r_{jt} = \mathbf{x}'_{jt}\beta^r_t - \alpha^r_t p_{jt} + \xi^r_{jt}.$$

[23]Note that we also use data from products with zero sales in a month in the estimation, but do not show them in the plot.

For each market $t$, we observe the true recommendation status of alternative $j$. If we ignore the unobserved quality component $\xi^r_{jt}$, we can form the log-likelihood

$$\mathcal{L}(\alpha^r, \beta^r) = \sum_t \sum_j \Big( \mathbf{1}\{j \text{ recommended on } t\} \times \log r_{jt}(\alpha^r, \beta^r) \Big).$$

However, even when prices and quality are uncorrelated, simply omitting $\xi^r_{jt}$ from the model will lead to inconsistent estimates. Hence, we want to model unobserved quality as a random effect. To do so, we change notation to $t = (p, \tau)$ and explicitly acknowledge the panel structure of the data. In our new notation, $p$ refers to a market and $\tau$ a date. We will then assume that

$$\delta^r_{jp\tau} = \mathbf{x}'_{jp\tau}\beta^r_t - \alpha^r_t p_{jp\tau} + \xi^r_{jp}, \text{ and } \xi^r_{jp} \sim N(0, \sigma^2_{\xi^r}),$$

where $\sigma_{\xi^r}$ is a parameter to be estimated. This gives (infeasible) log likelihood

$$\mathcal{L}(\alpha^r, \gamma^r, \sigma_{\xi^r}) = \ln \left\{ \prod_p \int \left[ \prod_\tau \prod_j r_{jp\tau}^{\mathbf{1}\{j \text{ recommended on } t\}}(\alpha^r, \beta^r, \sigma_{\xi^r} z_{jp}) \right] d\Phi(z_{jp}) \right\}.$$

We form a feasible approximation to this likelihood by employing simulation draws $s = 1, \ldots, S$ where $z_{jps} \sim N(0, 1)$:

$$\hat{\mathcal{L}}(\alpha^r, \gamma^r, \sigma_{\xi^r}) = \ln \left\{ \prod_p \frac{1}{S} \sum_s \left[ \prod_\tau \prod_j r_{jp\tau}^{\mathbf{1}\{j \text{ recommended on } t\}}(\alpha^r, \beta^r, \sigma_{\xi^r} z_{jps}) \right] \right\}.$$

In practice, we draw $(z_{jps})$ once before the maximization procedure is run. To reduce the risk of simulation error, we increase $S$ until the estimates stabilize.

## D.3 MSL Estimation of Consumer Choice

While exactly one alternative is recommended for each market $(p, \tau)$, our sales data is generated by the arrival of some (generally unknown) number of customers. These customers either buy one of the alternatives or select the outside option. We discussed above in Section D.1 how we estimate the market size $M_p$, i.e., the number of customers that choose between inside and outside option on any given date $\tau$ and a fixed market $p$. Given this number, we can form a pseudo-likelihood function[24] by noting that the number of actual sales will take a Binomial distribution. Write $a_{jp\tau}$ for the observed number of sales for alternative $j$. Since we only observe sales for one merchant,

[24] This is only a pseudo-likelihood as our predicted number of arrivals $M_p$ need not be an integer. In practice, this causes no issues as we can employ a continuous extension of the Binomial function when evaluating the likelihood.

we have

$$
\begin{aligned}
\mathcal{L}(\alpha,\beta,\sigma_\xi) &= \ln \prod_p \frac{1}{S} \sum_s \prod_\tau \prod_j \binom{\hat{M}_j}{a_{j^o(p),p,\tau}} \\
&\quad \times s_{j^o(p),p,\tau}^{a_{j^o(p),p,\tau}}(\alpha,\beta,\sigma_\xi,\hat{\alpha}^r,\hat{\beta}^r,\hat{\sigma}_{\xi^r})[1-s_{j^o(p),p,\tau}^{a_{j^o(p),p,\tau}}(\alpha,\beta,\sigma_\xi,\hat{\alpha}^r,\hat{\beta}^r,\hat{\sigma}_{\xi^r})]^{\hat{M}_j - a_{j^o(p),p,\tau}}.
\end{aligned}
$$

This likelihood features several significant departures from the 'usual' logit likelihood, which we now discuss in turn. First of all, as noted before, we proxy for the market size using sales ranks. Secondly, as our products are on the long tail, consumers arrive infrequently. Hence, it is likely untrue that the fraction of sales going to a merchant is a good proxy for purchase probability. Instead, since sales are distributed binomially with the provided purchase probability, we can form our likelihood on this basis.

Thirdly, our likelihood needs to consider the dependence of consumer choice on recommendation status through the consideration set model, as we discussed in the main text. We observe the recommendation status for each market whenever one of the alternatives on the market changes. However, the recommendation status can change without us seeing that it has changed. This fact leaves us with two options. We can take the last observed recommendation status and use this as a proxy for the current status. This proxy will introduce measurement error but has the benefit of being model-free. That is, our estimation will still be consistent even if our model of the recommendation algorithm is misspecified. Alternatively, we can rely on prior estimates of the parameters of the recommendation algorithm $(\hat{\alpha}^r,\hat{\beta}^r,\hat{\sigma}_{\xi^r})$. To the extent that our model is correctly specified, we gain robustness to measurement error from the recommendation status. However, we will introduce error in sampling variation when we estimate the parameters of the recommendation algorithm.[25] In practice, we ran the estimation procedure both ways and found only minor differences.

### D.4 Offer FE under Full Observability

We now illustrate how to condition out alternative-specific fixed effects when estimating a multinomial logit discrete choice model. We first consider the case when the chosen alternative is observed (as is the case for the estimation of the recommender system in the main text). Subsequently, we extend our results to the case where there is one alternative for each market, and we only observe whether or not this specific alternative is chosen. This section generalizes results in Chamberlain (1980) which themselves are based on Rasch (1960, 1961). See Arellano and Honoré (2001) for a modern treatment.

For this section, we assume that we have fixed some product page $p$ and suppress the relevant subscript. Extending the estimators from one to multiple product pages is trivial.

[25] We deem this second source of error to be less severe because we only need to estimate a few parameters with ample data.

Recall that the recommender system's mean utility from alternative $j$ at time $\tau$ is given by

$$\delta^r_{j\tau} = \mathbf{x}'_{j\tau}\beta^r - \alpha^r p_{j\tau} + \xi^r_j,$$

where in contrast to the main text we have imposed Assumption 7, i.e. offer quality $\xi^r_j \equiv \xi^r_{j\tau}$ is time invariant.

Mean utilities on each offer are combined with a Type-1 extreme value shock $\epsilon_{j\tau}$ to yield a utility index $v^r_{j\tau} = \delta^r_{j\tau} + \epsilon^r_{j\tau}$. The recommendation for market $t$ is assigned to the offer with the highest utility index. If $y^r_{j\tau}$ is a dummy for whether an alternative is recommended, we have

$$y^r_{j\tau} = \mathbf{1}\{v^r_{j\tau} \geq \max_k v^r_{k\tau}\}.$$

Using McFadden (1981), this multinomial logit model can be transformed into a standard binary logit model by appropriate conditioning. Indeed, consider two alternatives $j$ and $j'$ that have a non-zero probability of being recommended. Then,

$$\mathbb{P}\left(y^r_{j\tau} = 1 | y^r_{j\tau} + y^r_{j'\tau} = 1\right) = \frac{1}{1 + \exp\left(\delta^r_{j\tau} - \delta^r_{j'\tau}\right)}.$$

But as argued by Chamberlain (1980), a fixed effect in a binary logit model can be eliminated by (further) conditioning on an appropriate sufficient statistic. Fix two days $\tau, \tau'$ and consider the event $C = \{y^r_{j\tau} + y^r_{j'\tau} = 1, y^r_{j\tau'} + y^r_{j'\tau'} = 1\}$. Then

$$\begin{aligned}\mathbb{P}(y^r_{j\tau} = 1 | y^r_{j\tau} + y^r_{j\tau'} = 1, C) &= \frac{\mathbb{P}(y^r_{j\tau} = 1, y^r_{j\tau} + y^r_{j\tau'} = 1 | C)}{\mathbb{P}(y^r_{j\tau} + y^r_{j\tau'} = 1 | C)} \\ &= \frac{1}{1 + \exp[(\delta^r_{j\tau} - \delta^r_{j\tau'}) - (\delta^r_{j'\tau} - \delta^r_{j'\tau'})]}.\end{aligned}$$

But neither $\delta^r_{j\tau} - \delta^r_{j\tau'}$ nor $\delta^r_{j'\tau} - \delta^r_{j'\tau'}$ contain $\xi^r_j$. Thus, to consistently estimate $\beta^r$ and $\alpha^r$ in the presence of $\xi^r_j$ we can maximize the log likelihood function

$$\mathcal{L}(\alpha^r, \beta^r) = \sum_{j=1}^{|\mathcal{J}|} \sum_{\tau=1}^{T} \sum_{j', \tau' \in Z_{j\tau}} \ln\left(\frac{1}{1 + \exp[(\delta^r_{j\tau} - \delta^r_{j\tau'}) - (\delta^r_{j'\tau} - \delta^r_{j'\tau'})]}\right).$$

Here, $Z_{j\tau}$ is the set of all potential offers $j' \neq j$ and times $\tau' \neq \tau$ that satisfy $y^r_{j\tau} + y^r_{j\tau'} = 1$, $y^r_{j\tau} + y^r_{j'\tau} = 1$ and $y^r_{j\tau'} + y^r_{j'\tau'} = 1$. In practice, we estimate the model under the restriction $\beta^r = \mathbf{0}$ as there is very little variation in non-price offer characteristics.

### D.5 Offer FE under Partial Observability

For each product page, we observe sales for exactly one alternative $j$. This renders the identification strategy of D.4 toothless, because forming the required conditioning set $Z_{j\tau}$ requires observations on at least two alternatives. However, at the cost of some power, we can exploit the high-frequency nature of our data to construct an estimator that is consistent in the presence of arbitrary fixed effects. Recall that a consumer's mean utility is given by

$$\delta_{j\tau} = \mathbf{x}'_{j\tau}\beta - \alpha p_{j\tau} + \xi_j,$$

where in contrast to the main text we have imposed Assumption 7, i.e. that offer quality $\xi_j \equiv \xi_{j\tau}$ is time-invariant. Mean utilities on each offer are combined in the usual fashion with a Type-1 extreme value shock $\epsilon_{ij\tau}$ to yield a utility index $v_{ij\tau} = \delta_{j\tau} + \epsilon_{ij\tau}$. For now, assume all consumers are sophisticated, so each consumer simply chooses her preferred option. If $y_{ij\tau}$ is a dummy indicating whether consumer $i$ chooses alternative $j$ at time $\tau$,

$$y_{ij\tau} = \mathbf{1}\{v_{ij\tau} \geq \max_k v_{ik\tau}\}.$$

But note

$$\begin{aligned}
\mathbb{P}(y_{ij\tau} = 1 | y_{ij\tau} + y_{ij\tau'} = 1) &= \frac{\mathbb{P}(y_{ij\tau} = 1)\mathbb{P}(y_{ij\tau'} = 0)}{\mathbb{P}(y_{ij\tau} = 1)\mathbb{P}(y_{ij\tau'} = 0) + \mathbb{P}(y_{ij\tau} = 0)\mathbb{P}(y_{ij\tau'} = 1)} \\
&= \frac{1}{1 + \frac{\sum_{k \neq j} \exp(\delta_{ik\tau})}{\sum_{k \neq j} \exp(\delta_{ik\tau'})} \exp(\delta_{ij\tau'} - \delta_{ij\tau})}
\end{aligned}$$

In general, $\sum_{k \neq j} \exp(\delta_{ik\tau}) \neq \sum_{k \neq j} \exp(\delta_{ik\tau'})$. However, when $x_{k\tau} = x_{k\tau'}$ and $p_{k\tau} = p_{k\tau'}$ for all $k \neq j$, then equality holds. Thus, we restrict attention to pairs of days between which all offers for which we do not observe sales remain unchanged. This yields the following likelihood function:

$$\mathcal{L}(\alpha, \beta) = \sum_{(\tau, \tau') \in X} \ln \left( \frac{1}{1 + \exp(\delta_{ij\tau'} - \delta_{ij\tau})} \right),$$

where

$$X = \{(\tau, \tau') | \forall k \neq j : x_{k\tau} = x_{k\tau'}, p_{j\tau} = p_{j\tau'}\} \cap \{(\tau, \tau') | y_{ij\tau} = 1, y_{ij\tau'} = 0\}.$$

In practice, we estimate the model under the restriction $\beta = \mathbf{0}$ as there is very little variation in non-price offer characteristics.

### D.6 Across-Market Variation in Preference Parameters

As we discuss in the main text, our estimation imposes that the price coefficient in the recommendation system is specified as

$$\alpha^r_{p,\tau} = \frac{\overline{\alpha}^r}{R_p}.$$

A similar assumption was made for demand. Intuitively, we are restricting the way that the price coefficient can vary between various markets. In contrast to typical applications of discrete choice, our consumers are faced with distinct choice situations in different markets. In our context, a market is a product page. Thus, the way a typical consumer makes his choice between various offers will plausibly vary across these choice situations. Intuitively, minor price differences matter more for cheaper products. This is corroborated by the industry experts we interviewed.

We now investigate whether this assumption is appropriate for the recommender system.[26] To do so, we re-estimate column (3) of Table 5 for products with MSRP between \$0 and \$10, \$10 and \$20 and so on. The results are reported in Figure 13. Indeed, our specified functional form fits the data well.

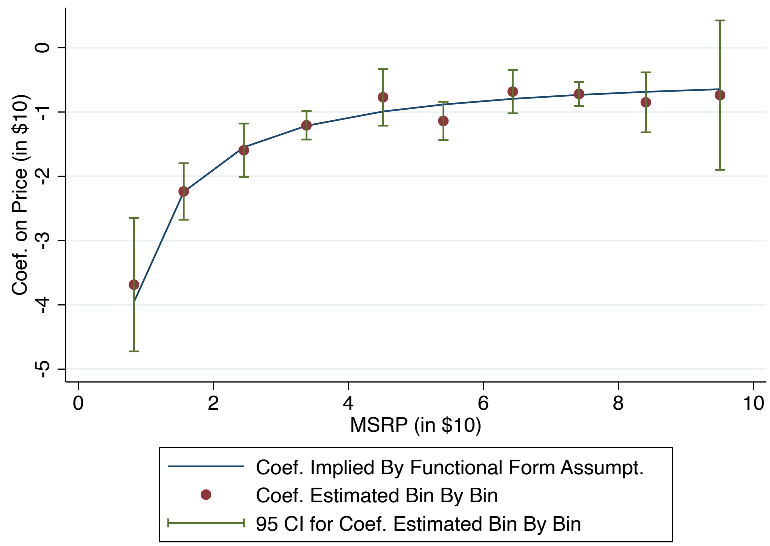



**Figure 13:** *How $\alpha^r$ Varies With MSRP.*

Notes*: This figure illustrates how our estimates of $\alpha^r$ vary across MSRP bins. The blue line shows the variation implied by our functional form assumption. The red dots illustrate the actual estimate for the coefficient on products just from this MSRP bin (with green CIs). Our specified functional form fits the data well.*

# E Details of Simulated Method of Moments

## E.1 Solving the Pricing Game

Fix a market $p$ and entrants $\mathcal{J}$. Each entrant's type is $(c_j, q_j, q_j^r)$, where $q_j = x_j'\beta + \xi_j$ and $q_j^r = x_j'\beta^r + \xi_j^r$ measure the attractiveness of the entrant to consumers and the recommendation system respectively. Entrants are fully informed about each others' types and treat price as a

[26]Though we lack the power to ask the same question in the context of consumer choice, we note that one would expect the recommendation system to try and match the price sensitivity of consumers. Further, unreported results suggest that the log-likelihood improves in demand when imposing this functional form (as compared to requiring a simple linear coefficient on price).

strategic variable. Thus, the first-order condition associated with $j$'s choice of price is

$$\phi s_j(\mathbf{p}, \mathbf{q}) + [\phi p_j - c_j] \frac{\partial s_j}{\partial p_j} = 0. \tag{2}$$

An equilibrium of the pricing game comprises prices $\mathbf{p}$ satisfying Equation (2) for all entrants $j$. To ease notation, suppress the dependence of all variables on $\mathbf{q}$. Following Morrow and Skerlos (2011), we rewrite the previous equation as a *$\zeta$-markup equation*. First we decompose the Jacobian matrix of market shares:

$$\frac{\partial \mathbf{s}}{\partial \mathbf{p}'} = \Lambda(\mathbf{p}) - \Gamma(\mathbf{p}),$$

where $\Lambda(\cdot)$ contains the diagonal elements of the Jacobian matrix and $\Gamma(\cdot)$ contains the factors common to both the diagonal and off-diagonal elements. Now, the $\zeta$-markup equation is

$$\mathbf{p} = \phi^{-1}\mathbf{c} + \zeta(\mathbf{p}), \quad \zeta(\mathbf{p}) \triangleq \Lambda(\mathbf{p})^{-1}\Gamma(\mathbf{p})'(\phi\mathbf{p} - \mathbf{c}) - \Lambda(\mathbf{p})^{-1}\mathbf{s}(\mathbf{p}), \tag{3}$$

when $\Lambda(\mathbf{p})$ is nonsingular.

Under some technical conditions, Proposition 2.12 in Morrow and Skerlos (2010) tells us that the $\zeta(\mathbf{p})$ term can be expressed as

$$\zeta(\mathbf{p}) = \Omega(\mathbf{p})(\phi\mathbf{p} - \mathbf{c}) + \Big(I - \Omega(\mathbf{p})\Big)\eta(\mathbf{p}),$$

where $\Omega \triangleq \Lambda^{-1}\Gamma$ and $\eta(\mathbf{p})$ is the standard BLP markup term

$$\eta(\mathbf{p}) \triangleq -\left[\frac{\partial \mathbf{s}}{\partial \mathbf{p}'}\right]' \mathbf{s}(\mathbf{p}).$$

Having described Equation (3), we chain iterations based on the $\zeta$-markup and BLP-markup equations to find a fixed point. Iterating on the $\zeta$-markup equation improves convergence relative to iterating on the BLP-markup equation.[27]

In our model, we note that an equilibrium exists but is unlikely to be unique. With two types of consumers, the pricing game is no longer supermodular (with a dominant diagonal), so the equilibrium we find will depend on starting values. Our estimation procedure gives the entrant with the highest recommendation system attractiveness a lower starting price than the others. In robustness checks, we find that varying the "privileged" entrant does not change our qualitative findings.

[27] There are examples for which "iterating on the BLP-markup equation is not necessarily locally convergent, while iterating on the $\zeta$-markup equation is superlinearly locally convergent" (Morrow and Skerlos 2011, p. 329).

## E.2 Solving the Entry Game

The information set[28] of potential entrants $j \in \mathcal{N}$ is $\mathcal{I}_j = \{c_j, F\}$. As profits are strictly decreasing in own costs, firms play cutoff entry strategies, i.e. $\chi_j = 1\{c_j \leq c_j^*\}$. From the perspective of prospective entrant $j$, entry is profitable if and only if

$$\mathbb{E}_{\mathbf{q},c_{-j}}[\pi_j(c_j, q_j; \mathbf{c}_{-j}, \mathbf{q}_{-j}, \chi_{-j})] \geq F.$$

We focus on symmetric equilibria, i.e. equilibria where $c_j^* = c^*$ for all $j \in \mathcal{N}$. This $c^*$ must satisfy

$$\underbrace{\mathbb{E}_{\mathbf{q},c_{-j}}[\pi_j(c^*, q_j; \mathbf{c}_{-j}, \mathbf{q}_{-j}, \chi_{-j}^*)]}_{=:V(c^*)} = F \text{ where } \chi_k^*(c) = 1\{c_k \leq c^*\} \text{ for all } k \neq j.$$

We can understand the LHS of this equation as a function $V(c)$ in one parameter. For each candidate cutoff $c$, we can approximate $V(c)$ by replacing the expectation with an average across simulation draws, i.e.

$$\hat{V}(c) = \frac{1}{S} \sum_{s=1}^{S} \pi_j(c^*, q_j^s; \mathbf{c}_{-j}^s, \mathbf{q}_{-j}^s, \chi_{-j}^*).$$

We can then use standard root-finding techniques on $\hat{V}(c) = F$ to find $c^*$.

## E.3 Simulating a Market

We can now combine E.1 and E.2 to simulate market-level outcomes. Formally, we employ the algorithm presented in Figure 14b. We begin by drawing a scalar fixed cost $F$. Furthermore, we draw $S$ vectors of wholesale costs $\mathbf{c}_s \in \mathbb{R}^{|\mathcal{N}|}$, demand qualities $\mathbf{q}_s \in \mathbb{R}^{|\mathcal{N}|}$ and recommender qualities $\mathbf{q}_s^r \in \mathbb{R}^{|\mathcal{N}|}$. We employ the simulation draws and the algorithm of E.2 to obtain $c^*$. Next, we fix simulation draw $s = 1$ and find the set $\mathcal{J} = \{j \in \mathcal{N} : c_{js} \leq c^*\}$ of successful entrants. We then calculate their equilibrium prices, profits and market shares using E.1.

## E.4 Aggregating Markets to Moments

We will now aggregate our market-level outcomes to moments; our discussion follows Ackerberg (2009) adapted to our model. Fix a market $p$. Our model postulates a relationship $f$ between (market-level) observables[29] $x_p = (M_p, R_p)$, unobservables $u_p$ and outcomes $y_p$. In particular, at the true parameter vector $\theta_0$ we have

$$y_p = f(x_p, u_p, \theta_0).$$

[28] We suppress the fact that fixed cost also enter the information set here as we are already conditioning on a market and we assume that fixed costs do not vary within market.

[29] Here, $M$ refers to market size and $R$ to the manufacturer's suggested retail price.

We begin by defining the outcomes included in $y_p$. Let the number of successful entrants $J_p$, the average price $\overline{p}_p$ and the standard deviation of log prices $std_p$ on market $p$ be given by

$$J_p := |\mathcal{J}_p| = \sum_{j\in\mathcal{N}_p} 1\{c_{jp} \le c_p^*\}, \quad \overline{p}_p = \frac{1}{J_p}\sum_{j=1}^{J_p} p_{jp}, \text{ and } std_p = \sqrt{\frac{1}{J_p}\sum_{j=1}^{J_p}[\log(p_{jp}) - \overline{\log(p_p)}]^2}$$

The outcomes we want to match are $y_p = (y'_{p,1}, y'_{p,2})'$ where

$$y_{p,1} = \begin{pmatrix} J_p \\ J_p^2 \\ J_p \times A_p \\ J_p \times R_p \\ 1\{J_p = 1\} \\ \vdots \\ 1\{J_p = \overline{N}\} \end{pmatrix}, \qquad y_{p,2} = \begin{pmatrix} \overline{p}_p \\ \overline{p}_p \times R_p \\ std_p \end{pmatrix}.$$

If the data $\{x_p, y_p\}_{p=1}^P$ is generated by by our model at the true $\theta_0$, then

$$\theta = \theta_0 \implies \mathbb{E}[y_p - \mathbb{E}[f(x_p, u_p, \theta)|x_p]|x_p] = 0. \tag{4}$$

As long as our model parameters are identified (we argue they are in the main text), the reverse implication also holds. As econometricians, we do not observe the true value of the unobservables $u_p$. In our model, these shocks include (i) fixed costs $F_p$ for each market and (ii) qualities $(q_{jp}, q_{jp}^r)$ as well as wholesale unit costs $c_{jp}$ for each potential entrant on each market. To proceed, we make parametric assumptions on the distributions of these quantities. Concretely, we assume that $(q_{jp}, q_{jp}^r)$ are drawn at random from the empirical distribution implied by our recommender system and consumer choice estimation. Fixed costs are drawn from

$$F_p \sim \text{LogN}(\theta_0^F + x_p'\theta_x^F, \theta_\sigma^F).$$

and wholesale costs are drawn from

$$c_{jp} \sim \text{LogN}(\theta_0^c + x_{jp}'\theta_x^c,\ \theta_\sigma^c).$$

Given some candidate $\theta = (\theta_0^F, \theta_x^F, \theta_\sigma^F, \theta_0^c, \theta_x^c, \theta_\sigma^c)$, the conditional distribution of unobservables $p(u_p|x_p, \theta)$ is thus fully specified. In theory, we could use our model and the moment condition(s) in (4) to estimate $\theta$. However, in practice these expectations are hard to compute: they involve not one but two layers of games for which equilibria must be numerically computed (the entry game and the pricing game). Hence, we instead take simulation draws $(u_{p1}, \ldots, u_{pS}) \sim p(u_p|x_p, \theta)$ and approximate the expectation by averaging:

$$\widehat{\mathbb{E}f_p(\theta)} = \frac{1}{S}\sum_s f(x_p, u_{ps}, \theta).$$

As shown in McFadden (1989) and Pakes and Pollard (1989), estimation can proceed based on the simulated method of moments estimator that sets the simulated moment vector

$$\hat{G}(\theta) = \frac{1}{P}\sum_p (y_i - \widehat{\mathbb{E}f_p(\theta)})$$

as close as possible to zero. Concretely, our estimator will solve

$$\hat{\theta} = \arg\min_{\theta\in\Theta} Q(\theta) = \hat{G}(\theta)'W\hat{G}(\theta)$$

for some weight matrix $W$.

### E.5 Concentrating Out Wholesale Cost Parameters

We can partition $\theta = (\theta^F, \theta^c)$ where $\theta^F = (\theta_0^F, \theta_x^F, \theta_\sigma^F)$ and $\theta^c = (\theta_0^c, \theta_x^c, \theta_\sigma^c)$. Define

$$\tilde{\theta}^F(\theta^c) := \arg\min_{\theta_F} Q_n((\theta^F, \theta^c)), \quad \text{and} \quad \tilde{\theta}^c := \arg\min_{\theta^c} Q_n((\theta^F(\theta^c), \theta^c)).$$

Then it is easy to see that $\hat{\theta} = (\tilde{\theta}^F(\tilde{\theta}^c), \tilde{\theta}^c)$. Hence, we can 'concentrate out' the fixed cost parameters when searching over possible values of the unit cost parameters. Formally, we employ the algorithm presented in Figure 14a.

Our approach is beneficial for the same reason the linear parameters are concentrated out in the estimation of Berry, Levinsohn, and Pakes (1995): there is little computational burden in estimating the parameters that were concentrated out. In BLP, the linear parameters can be estimated in closed form. For us, $\tilde{\theta}^F(\tilde{\theta}^c)$ can be found without re-solving the pricing games.

### E.6 Importance Sampling

Computationally, finding $\tilde{\theta}^F(\theta^c)$ requires a continuous objective. However, the entry game is discrete: e.g., for some markets, as we increase $F$, there may be exactly one entrant until we hit some threshold, after which there are two. To avoid these discontinuities, we employ an importance sampling technique advocated by Ackerberg (2009). Recall

$$\ln F_p = \theta_0^F + x_p'\theta_x^F + u_p^F.$$

As $\theta^c$ is fixed, we will omit the dependence of outcomes on it in our notation. Thus, we can write outcomes directly as a function of fixed cost draws rather than as a function of shocks, i.e. $f(x_p, u_p, \theta) = \tilde{f}(F_p)$. This suggests the following simulation strategy: draw $F_{ps} \sim g(\cdot)$ where $g(\cdot)$

**for** $p = 1, \ldots, P$ **do**
  **for** $s = 1, \ldots, S_F$ **do**
    simulate entry for (p,s)
$\theta^F = \theta^F_0$
**while** *dist* > *tol* **do**
  reweight sim data for $\theta^F$
  dist = $\hat{G}(\theta^c, \theta^F)$
  update $\theta^F$
**return** $\hat{G}(\theta^c, \theta^F)$

**(a)** *SMM Moments* $\hat{G}(\theta^c)$.

**Data:** $S_E$ sets $\mathcal{N}_{s_e}$ of pot entrants
**for** $\tilde{c} \in grid$ **do**
  fix $c_1 = \tilde{c}$
  **for** $s_e = 1, \ldots, S_E$ **do**
    $\mathcal{J} = \{1\} \cup \{j \in \mathcal{N}_{s_e} : c_j < \tilde{c}, j > 1\}$
    find $\pi_1(\mathcal{J})$ using fix-point iter.
  $V(\tilde{c}) \approx \frac{1}{S_E} \sum_{s_e} \pi_1(\cdot)$
$c^*$ solves $V(c^*) = F$
$\mathcal{J} = \{j \in \mathcal{N}_1 : c_j < c^*\}$
solve pricing game given $\mathcal{J}$

**(b)** *Simulation of Entry.*

**Figure 14:** *Algorithms Used in SMM Estimation.*
Notes*: We provide the two key algorithms used in the SMM estimation procedure. In the left panel, we detail how to find the value of the* outer *SMM objective. In the right panel, we describe in more detail the key step of entry simulation.*

is a heavy-tailed density. Then use our knowledge of the distribution of fixed costs implied by our current parameter guess to re-weight the outcomes at the simulated fixed cost draws. Formally speaking,

$$\widetilde{\mathbb{E}f_p}(\theta) = \sum_s \tilde{f}(F_{ps}) \frac{p(F_{ps}|x_p, \theta)}{g(F_{ps}|x_p)}.$$

This simulator has two advantages. The primary advantage in our case is the fact that the density for fixed costs that we specified is smooth, which ensures that the simulated outcomes will smoothly depend on $\theta^F$. So, for instance, as we increase the mean of the fixed cost distribution, the estimator will smoothly put less and less weight on low fixed cost draws.

The second advantage is that we only need to compute the market-level outcomes $\tilde{f}(F_{ps})$ once at the beginning of the procedure. As we vary $\theta^F$, the outcomes employed do not vary but are instead simply reweighted. The computational savings from this are usually substantial but are less pronounced here.

Crucially, following this importance sampling procedure could introduce extra simulation error into our estimates. For this reason, we follow the recommendation in Ackerberg (2009) and iterate the procedure several times: at each new iteration, we let $g(\cdot|x_p)$ equal the density $p(\cdot|x_p, \theta^F_{prev})$ at the previous' iterations optimal value $\theta^F_{prev}$. As in the original paper, we find that this iteration converges extremely fast (typically in just three steps).

### E.7 Results

We illustrate the fit of our SMM estimates in Figure 15. The top row compares the model distributions of product-level outcomes implied by the estimated $\hat{\theta}$ (outlined in blue) to the distribution of these quantities in the data. In particular, Figure 15a illustrates the empirical distribution of mean price across products (in solid light orange) against the distribution implied by the model (outlined in blue). We see that the model distribution matches that in the data extremely well. Similarly, Figure 15b illustrates the distributions of the number of entrants. The fit is also good with one key exception: the model has a hard time rationalizing the excess mass at the point where we censor the distribution of the number of entrants: a surprisingly large number of products have more than 15 entrants in the data. We suspect that some of these products may have 'less serious' entrants, such as drop shippers who will try to fulfill an order by immediately ordering the same product on a different e-commerce platform (and entering their customer's address as the delivery address).

Moving to the bottom row, we can evaluate to what extent our model is capturing across-product heterogeneity. For example, in 15c, we can observe that products with high prices in the data also have high prices in our model. This is due primarily to our exploitation of MSRP data, an excellent proxy for costs. By contrast, 15d illustrates that our model cannot capture much of the across-product variation in the number of entrants. This will preclude us from making statements about e.g. potentially heterogeneous effects of the design of the recommendation algorithm.

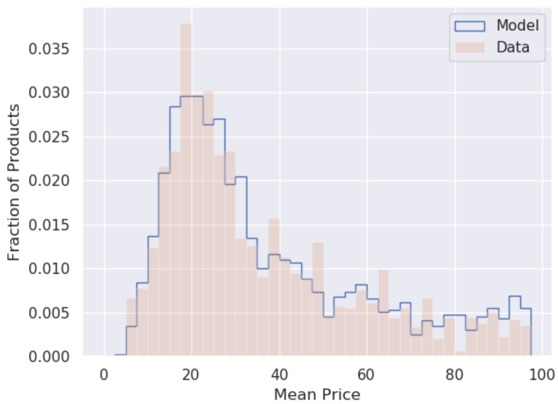
Why can our model not match the across-product variation in the number of entrants? On the one hand, we may just be missing good covariates that determine fixed costs. For instance, we have allowed fixed costs to vary by market size and suggested retail price. Perhaps, however, fixed cost varies across brands: this could occur if some companies refuse to sign deals that they consider 'peanuts' while other companies are happy to supply their product at whatever quantity a retailer requests. As econometricians, we do not observe this heterogeneity.

This reasoning naturally suggests a plan of attack. We could exploit the (excellent) information we do have on fixed costs: each product's observed number of entrants. However, this approach is dangerous. What if there is just random variation in (effective) fixed costs? Concretely, it could be that discovering niches to enter is just a fundamentally random process. As already discussed in the main text, turnover on Amazon is high. This suggests that there are a lot of unmodelled shocks to entry decisions. Under this line of argument, attempting to 'force' all dots in Figure 15d to lie on the 45-degree-line would amount to overfitting. Hence, we take the more cautious approach of restricting ourselves to examining aggregate effects.

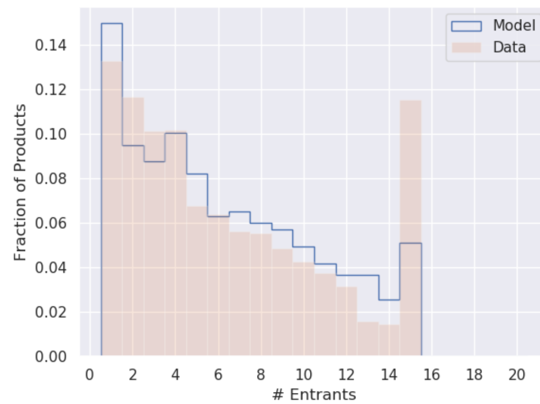
# F Additional Counterfactuals

## F.1 Platform Price Response to Self-Preferencing

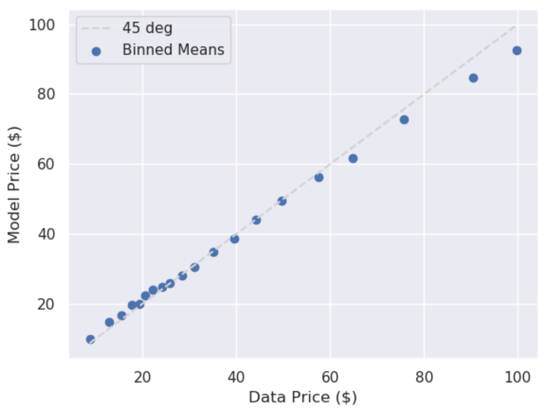
Our mainline results on the effect of self-preferencing do not account for the platform's pricing strategy. Indeed, we assume in the main text that the platform's prices and entry decisions remain fixed across all counterfactuals. Here, we relax this assumption. Rather than take a stance on the objective function that the platform is optimizing, however, we instead compute the price change required for our welfare results to flip. To do so, we progressively lower Amazon's prices in the counterfactual world without self-preferencing, continuing until we have reached the point at which the overall welfare effect of this platform advantage is (close to) zero. This point is reached



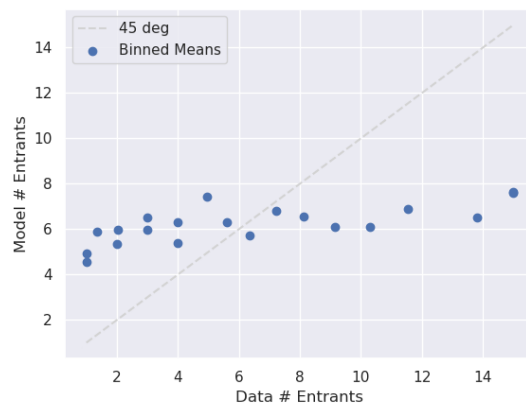

**(a)** *Distribution of Mean Price.*




**(b)** *Distribution of # Entrants.*




**(c)** *Fit of Mean Price.*




**(d)** *Fit of # Entrants.*

**Figure 15:** *Fit of Distributions Implied By $\hat{\theta}$.*

Notes: *This figure provides the distributions implied by the values of $\theta$ estimated as part of the simulated method of moments procedure and compares them to the empirical distributions.*

once Amazon's prices are 7.8% higher in the observed scenario, relative to the counterfactual world. To be exact, we exhibit the overall effect of self-preferencing under this pricing difference in Table 14.

| Δ Outcome/Counterfactual | Static | Pricing only | Full (Pricing + Entry) |
|---|---|---|---|
| Δ Platform Fees | \$10,142 | \$9,683 | \$14,438 |
| Δ Total Consumer Surplus | -\$67,439 | -\$62,818 | -\$50,947 |
| Δ Consumer Surplus (Naive) | \$161,039 | \$228,415 | \$219,382 |
| Δ Consumer Surplus (Soph.) | -\$228,478 | -\$291,233 | -\$270,329 |
| Δ Producer Surplus | \$18,785 | \$50,935 | \$51,913 |
| Δ Welfare | -\$48,654 | -\$11,884 | \$966 |
| Δ Mean (Price/MSRP) | 0.00% | 0.26% | 0.20% |
| Δ Mean # Entrants | 0.00 | 0.00 | 0.02 |
| Δ Mean # Quality | 0.00 | 0.00 | -0.00 |
| Δ Mean # Sales/Month | 0.02 | 0.00 | -0.00 |

**Table 14:** *Self-Preferencing is Welfare Neutral if it Raises Platform Prices 7.8%.*

Notes*: This table displays differences in outcome variables when the owner of the marketplace advantages its own offers in its recommender system and raises prices by 7.8%, relative to a counterfactual in which it does not. The "Static" scenario assumes that sellers cannot change their prices or their entry decisions. The "Pricing only" scenario allows sellers to set prices optimally, but does not permit them to change their entry decisions. Finally, the "Full" counterfactual allows sellers to change both their prices and their entry decisions. Note that the overall welfare-effect of self-preferencing, \$996, is essentially zero — this is not a result but rather the way this table was derived: by searching for the amount by which the platform would have had to raise prices due to self-preferencing for the overall welfare effect of self-preferencing to be zero.*

In conclusion, the welfare effect of self-preferencing crucially hinges on whether one believes it is plausible for Amazon to have raised prices by more than 7.8% relative to the counterfactual world in which no self-preferencing occurs.

### F.2 Making Recommendations More Price Elastic

In the main text, we estimate the value of search guidance relative to a counterfactual in which recommendations are made uniformly at random. However, we may also be interested in the opposite, i.e. how the estimated algorithm performs relative to an algorithm that is even more aggressive in its emphasis on price. We investigate this question by evaluating a counterfactual algorithm which doubles the estimated price coefficient, and reporting in Table 15 the impact of moving from such a counterfactual algorithm to the factual algorithm actually employed.

We find that, compared to a heavier emphasis on price, the algorithm employed in practice performs better. To begin with, the first column of Table 15 indicates that before taking pricing and entry decisions into account, the estimated algorithm benefits consumers, producers and platform alike: its comparatively weaker (but still strong) emphasis on price allows the platform to surface higher quality offers to consumers, which these are more likely to purchase. Once we allow pricing to change, the weaker emphasis on price leads to the expected increase in prices. However, at 0.61%, the effect on prices is quite small. This is because in the more aggressive pricing scenario, some merchants abandon their attempts to capture the recommendation: giving

| Δ Outcome/Counterfactual | Static | Pricing only | Full (Pricing + Entry) |
|---|---|---|---|
| Δ Platform Fees | $45,230 | $55,369 | $47,520 |
| Δ Total Consumer Surplus | $328,521 | $75,237 | $405,929 |
| Δ Consumer Surplus (Naive) | $328,521 | -$96 | $69,962 |
| Δ Consumer Surplus (Soph.) | $0 | $75,333 | $335,967 |
| Δ Producer Surplus | $77,535 | $130,147 | -$65,348 |
| Δ Welfare | $406,056 | $205,384 | $340,581 |
| Δ Mean (Price/MSRP) | 0.00% | 0.61% | -0.81% |
| Δ Mean # Entrants | 0.00 | 0.00 | 0.02 |
| Δ Mean # Quality | 0.00 | 0.00 | -0.00 |
| Δ Mean # Sales/Month | 0.11 | 0.08 | 0.06 |

**Table 15:** *The Estimated Algorithm Performs Better than a More Price-Elastic Variant.*
Notes*: This table displays differences in outcome variables when a recommender system is employed on our third-party marketplace, relative to a counterfactual in which the platform doubles its sensitivity to price. The "Static" scenario assumes that sellers cannot change their prices or their entry decisions. The "Pricing only" scenario allows sellers to set prices optimally, but does not permit them to change their entry decisions. Finally, the "Full" counterfactual allows sellers to change both their prices and their entry decisions.*

up on naive consumers can be a best response if targeting is too costly. Finally, there is slightly more entry under the factual algorithm, though, in line with the small price effect, the difference of 0.02 entrants per market is small.