Two-Sided Markets Shaped By Platform-Guided Search

Kwok Hao Lee, Leon Musolff

September 11, 2024

Abstract

Consider firms that operate platforms matching buyers and sellers while selling goods themselves. By guiding consumers towards their own products through algorithmic recommendations, these firms could influence market outcomes – a regulatory concern. To investigate, we combine novel data about sales and recommendations on Amazon with a structural model that captures seller entry. Recommendations are highly price-elastic (-20), and many consumers (34%) only consider recommended offers. Hence, algorithmic recommendations raise the demand elasticity (from -8 to -11), intensify price competition, and increase the purchase rate. However, increased competition reduces entry (but the missing merchants are the least efficient). Focusing on self-preferencing: recommendations favor Amazon (equivalent to a 6% price discount), but this skew does not act as a barrier to entry or otherwise harm consumers. Indeed, since consumers prefer Amazon's offers, "self-preferencing" slightly raises consumer surplus by $9 per product per month (assuming Amazon's prices remain constant.)

Keywords: Two-Sided Markets, E-Commerce, Consumer Search, Entry

[1]Lee: Department of Strategy and Policy, National University of Singapore Business School, kwokhao@nus.edu.sg; Musolff (corresponding author): Business Economics and Public Policy Department, The Wharton School, The University of Pennsylvania, lmusolff@wharton.upenn.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, Giovanni Compiani, Chris Conlon, Michael Dinerstein, Glenn Ellison, Qiang Fu, Jean-François Houde, Ian Kaneshiro, Uri Miron, Aviv Nevo, Devesh Raval, Tiffany Tsai, Raluca Ursu, and Matthijs Wildenbeest.

1 Introduction

This paper examines the interplay between Amazon's need to attract merchants and its ability to guide consumer search. The platform faces a tradeoff between fostering competition and incentivizing entry. This tradeoff is complicated by the company's participation on its platform as a seller. Because profits from sales serve as 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).[2] This concern has attracted attention from antitrust agencies (medianama.com) 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: To what extent can marketplace owners influence market outcomes in their favor? Rigorous empirical analysis has lagged popular attention for two reasons. First, scant data tie recommendations directly to demand. Second, evaluating harm from reduced entry has been stymied by computational intractability: entry and pricing decisions must be solved for repeatedly, ballooning the search time for supply parameters.

We 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 (e.g., Seim 2006). Entry is made tractable by harnessing frontier optimization routines and algorithmic improvements.[3]

Secondly, we combine our model with proprietary data that, for the first

[2]The usual definition of self-preferencing requires demonstrated harm to consumers, e.g., if Amazon favors itself in its recommendation algorithm *and, therefore,* consumer welfare is reduced. We are careful to use quotation marks ("self-preferencing") to refer to the first half of that statement, i.e., Amazon favors itself when recommending offers to consumers, given that our results suggest that Amazon is *not* self-preferencing (under the usual definition).

[3]These algorithmic improvements span Nash-Bertrand pricing iterations (Morrow and Skerlos 2011), importance sampling (Ackerberg 2009), and a new derivative-free optimization algorithm for determinants of fixed and wholesale costs (Cartis et al. 2019).

time, directly links sales and search guidance decisions on Amazon, a setting of prime regulatory concern. Since we essentially observe the "ground truth" about who is recommended *when* a sale is made, we can recover determinants of recommendations and demand. Still, estimation is challenging: we account for the potential endogeneity of prices *and* recommendations while utilizing information from just one alternative per market as well as proxies of total market sales (building on Chevalier and Mayzlin 2006). Combined with the entry model, we speak directly to the dynamic theory of harm from self-preferencing and the tradeoff between attracting merchants and fostering competition.

Our paper builds on the observation that platforms commonly set defaults for consumers through their search, ranking (Ursu 2018), and recommendation algorithms. Economists have increasingly become aware that these defaults can influence consumer choice (Thaler and Sunstein 2008).[4] We thus model consumers as having one of two kinds of consideration sets (Goeree 2008): either they consider all available options, or they only consider the recommended offer and the outside option. Through its influence on consideration sets, the platform's recommendation algorithm affects seller pricing (as in, e.g., Dinerstein et al. 2018), as well as *how many* and *which* firms enter.

We estimate our model on high-frequency data from Amazon, a critical example of a firm that participates as a merchant on its own e-commerce platform. Indeed, we find that the firm has intermediation power: 34% of consumers only consider offers recommended by the platform, thereby allowing the algorithm to raise the average price elasticity of demand from 8 to 11. Furthermore, it uses this power to intensify price competition (recommendations load heavily on price with an elasticity of 20) and preferentially guide consumers towards its own offers (which have an advantage equivalent to a 6% price discount).

After recovering the recommendation, demand, and cost parameters, we simulate market outcomes under two 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 platform companies' results when running brief experiments ("A/B tests") on their algorithms, as is standard industry practice. Secondly, in the "medium-run", we allow merchants to adjust their pricing decisions. Finally,

[4] As have firms: in 2018, Google paid Apple $12B to be Safari's default search engine (yahoo.com).

our innovation on the literature lies in our modeling of entry decisions, which we allow to vary in the "long-run" counterfactuals.

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, "self-preferencing" slightly raises consumer welfare because consumers prefer Amazon's offers. In the medium and long run, we find a negligible impact of this alleged anticompetitive practice on prices and entrants. Overall, the platform's recommendation advantage increases consumer welfare by $0.7 billion, or $9 per product per month. A supplemental analysis shows that if Amazon raised prices by more than 2.0%, "self-preferencing" could harm consumers. Still, our model does not validate the theory that "self-preferencing" harms consumer welfare by acting as a barrier to entry.

More stark is the positive overall effect of search guidance on welfare. Consumers enjoy a static welfare gain of $30 billion from Amazon's recommendation algorithm. Intuitively, offers on online marketplaces can vary dramatically, and guiding consumers towards attractive offers is very helpful. However, these gains rely on the recommendation algorithm's emphasis on price. Allowing merchants to adjust their prices, we find that the recommender system decreases the price of the cheapest offer on a product by, on average, 28% of the Manufacturer's Suggested Retail Price (MSRP). 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 $84 billion in the medium run). However, this redistribution is not without consequences. In the long run, the presence of recommendations causes a reduction in entry of 4.4 entrants/market. These "missing merchants" are primarily on the competitive fringe: with the recommender system, average wholesale costs decline by 17% of MSRP. Nonetheless, the entry channel constrains the platform from driving prices down further for fear of the "best" sellers exiting.

The paper proceeds as follows. Section 2 describes the literature. Section 3 introduces the Buybox algorithm on the Amazon Marketplace. To analyze how Amazon's recommender system impacts market outcomes, we formulate a model in Section 4 that we take to data, as described in Section 5. Section 6 displays results from our counterfactual analysis. We conclude in Section 7.

2 Literature Review

This paper speaks to three 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 (De Los Santos, Hortaçsu, and Wildenbeest 2012; 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), sometimes even favoring themselves over other sellers (Aguiar, Waldfogel, and Waldfogel 2021). 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). 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 entry.

In particular, the Amazon marketplace, and specifically its Buybox algorithm, has been a prime setting for much recent theoretical (e.g., Ciotti and Madio 2023) and empirical work on price competition and search on platform marketplaces (Chen and Tsai 2023; Crawford et al. 2022; Gutiérrez 2021; Lam 2021; Raval 2022). See Etro (2022) for a survey. To this literature we make three contributions. Our key empirical contribution is to assess the plausibility of Amazon's "self-preferencing" through the Buybox, as well as the impact of recommender systems more generally on merchant entry, pricing, and consumer demand. To this end, we formulate a new model of entry and Bertrand-Nash competition that features consumers with limited consideration sets. Finally, this model is taken to unique data that, for the first time, allows us to link recommendation status with sales on the platform.

Secondly, while the interplay between pricing and entry is novel to the platform-guided search literature, a literature on the value of variety appreciates the importance of additional draws from the entrant distribution. Online markets benefit from essentially "infinite shelf space", raising product variety in many markets

(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 2023). However, taste heterogeneity is not needed for additional sellers to benefit consumers; marginal entrants may be of high quality when such quality is ex-ante hard to predict (Aguiar and Waldfogel 2018).[5] We build on this literature by considering the value of merchants that would have entered Amazon marketplace had it not been for the intense price competition induced by the recommender system. Our results suggest that there is room for the platform to intensify price competition further to screen out less efficient merchants. The value of the "long tail" lies in more products rather than more offers on the same product.

Thirdly, we build on the literature on estimating entry costs (Bresnahan and Reiss 1991; see Berry and Reiss 2007 for a survey) and selective entry into auctions (see Hendricks and Porter 2007 for a survey, and Bodoh-Creed, Boehnke, and Hickman 2021 for an example on the eBay marketplace). In the latter series of 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. We follow Roberts and Sweeting (2010, 2013) and make an intermediate informational assumption: each potential entrant receives a signal about her value before she decides whether to enter. Although our simulation-based estimator for costs levies a heavy computational burden, our entry model remains tractable, and our equilibria can be characterized using wholesale cost cutoffs. An agent enters if and only if her signal says that entry will yield positive expected profits.

3 Who does the Amazon Buybox recommend?

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 utilized this opportunity (aboutamazon.com). Their listings were responsible for 60% of retail sales, yielding estimated merchant profits of $25 billion (aboutamazon.com).

[5]For instance, after the General Data Protection Regulation came into force, consumer surplus was decreased through this channel as fewer new apps entered the Google Play Store (Janssen et al. 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 (abouta-mazon.com). This speed suggests that search, ranking, and recommendation algorithms play an outsized role in shaping consumers' choices.

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

When multiple offers are present, the platform automatically designates *at most one*[6] 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 (europa.eu). Indeed, "industry experts estimate that about 80% of Amazon sales go through the Buy Box" (U.S. Senate Judiciary Committee 2020).

For econometricians, this algorithm offers an ideal setting to study 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 choose between various offers. For instance, the merchant "Zappos" offers these shoes for US$57.68 and will deliver them via Fulfillment by Amazon (FBA).[7] Alternatively, "BHFO" charges only US$54.99 but does not offer Fulfillment by Amazon. Below, we 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 our 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

[6] In rare situations, two offers are recommended: one Prime offer, and one non-Prime offer.

[7] Fulfillment by Amazon "is a fulfillment service that allows businesses to use Amazon to store, pick, pack, and ship customer orders."

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 competing on products that Amazon sells can be challenging 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).

There have also been recent allegations about the company modifying its search algorithms (wsj.com), favoring products from its own retail unit (Farronato, Fradkin, and MacKay 2023). 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" (thehill.com). Nevertheless, antitrust authorities worldwide continue to investigate this issue (europa.eu).

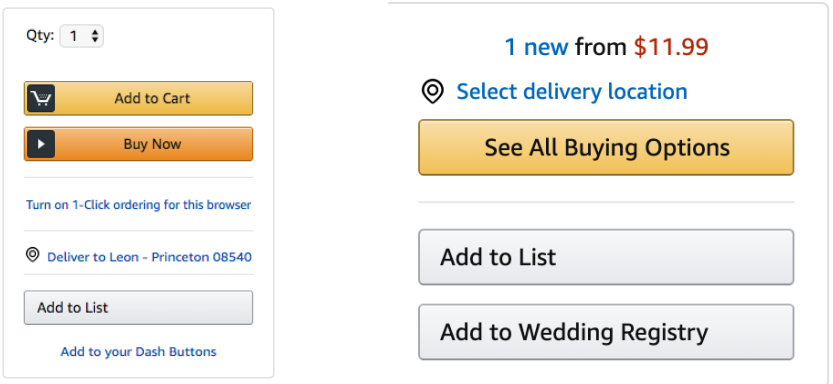
3.1 Data

We procured extensive, high-frequency data that is novel to this literature. Our dataset includes prices, recommendations, and sales on 50,486 products sold by 49,069 distinct merchants from 08/26/2018 until 03/25/2020. We provide summary statistics and discuss our data in more detail in Appendix A.

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. For each of the products 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 notified if the only change on a given product

is which offer was recommended.

We further observe, via a different API, all sales (and their exact time) for *a single merchant* in each market — the one using the repricing company's services. These data come with their own challenges (such as having to proxy for the market size), which we address with our model and estimation strategy in Sections 4 and 5. Yet, to our knowledge, no past papers have employed sales data across many merchants (one per market) to credibly estimate demand on Amazon Marketplace. In the absence of such data, employing sales ranks as a proxy has become common. However, one drawback of using sales ranks is that they aggregate sales at the product-page level. Hence, they do not allow the econometrician to uncover the relationship between the recommendation status of an offer and its market share.



(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), no seller has been recommended.

3.2 Canceled Recommendations

As a warmup, we investigate the platform's incentives to make *any* recommendation. When the platform does not recommend any offer, the Buybox is left empty, as in Figure 1b. While an empty Buybox likely lowers sales, the threat of canceled recommendations can be used to discipline prices: sellers are prompted to lower their prices to be recommended by the algorithm. Indeed, before 2019, Amazon's US marketplace required third-party sellers on its platform to sell their products

for a lower price than on other platforms (theverge.com). These "most-favored nations clauses" (MFN) were found to be anticompetitive by European antitrust agencies, leading Amazon to drop them in those markets by 2013.[8] Following similar threats, Amazon also dropped the MFN clauses for its US sellers by March 2019.

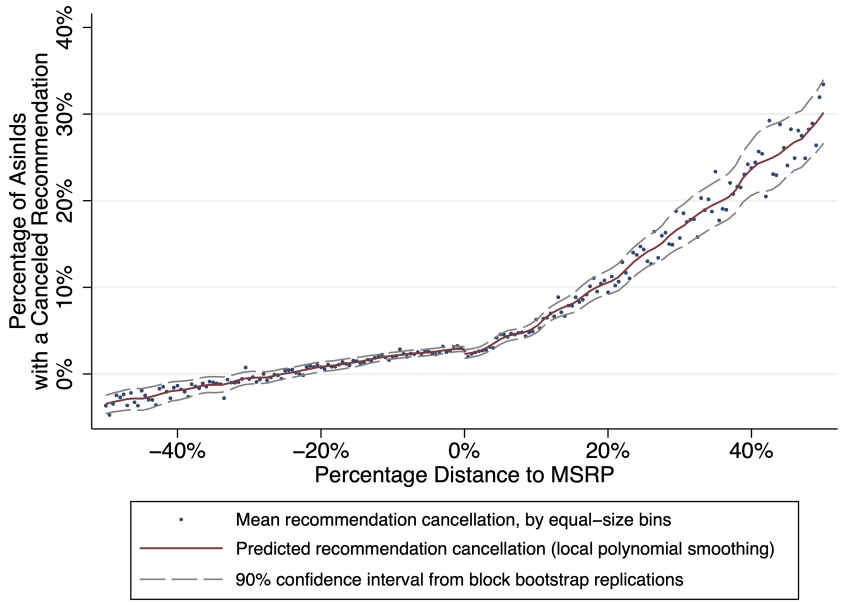


Figure 2: Evidence on Canceled Recommendations (Binscatter & Lowess).
Notes: This figure illustrates the relationship between canceled recommendations and uncompetitively priced offers. As the cheapest offer's markup over MSRP increases (moving from left to right), canceled recommendations become increasingly likely. More formally, let p index products and τ dates. We regress the cancellation fraction $\text{canceled}_{p\tau}$ for product p on date τ on the log difference between the price of the cheapest offer $\ell_{p\tau}$ and the MSRP for the product $R_{p\tau}$ (separately for prices larger and smaller than MSRP) while controlling for product fixed effects α_p. The plot shows the estimated relationship between $canceled_{p\tau} - \hat{\alpha_p}$ (on y) and $\log(\ell_{p\tau}) - \log(R_{p\tau})$ (on x.) All standard errors are clustered at the product level.

Nonetheless, canceled recommendations may achieve a similar effect as explicit MFN[9]. To demonstrate, we combine[10] data on the price of the lowest-priced offer

[8] 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 2021).

[9] This issue is currently subject to litigation in American courts (District of Columbia 2021)

[10] We collapse our data by product and date; for each product-date pair, we call the Buybox

on product p at date τ with data on Buybox ownership and information[11] on the manufacturer's suggested retail price (MSRP). Figure 2 illustrates the relationship between these variables after cleaning out product fixed-effects: 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. For offers priced above MSRP, a 1% increase in relative price is associated with a 0.55% higher probability that the recommendation will be canceled.

3.3 Recommendation Descriptives

If the platform decides to recommend one of the offers, how does it choose between them? We display some descriptive statistics. Figure 3a illustrates that there is within-product price variation; meanwhile, Figure 3b suggests that the within-product price rank is an important determinant of recommendation status. Though the cheapest offer is recommended almost 50% of the time, the second cheapest is recommended more than 25% of the time, and even more expensive offers take an (ever decreasing) share of recommendations.

Why is the lowest-priced offer not always recommended? Table 1 investigates possible determinants of recommendation status. For instance, the recommended offer is the cheapest in 41.63% of cases. However, offers Fulfilled by Amazon (FBA) are recommended in 86.52% of cases, perhaps because FBA offers dispatch much faster than others. This dispatch time advantage also accrues to Amazon. While Amazon's own offers are only recommended in 11.91% of observations, this is mostly because we observe more data on products without Amazon offers. For products featuring an Amazon offer (see the rightmost column of Table 1), the platform recommends its own offer 64.73% of the time. We verified in a prior version of this paper that, even among products with an Amazon offer, the same relationship between price and Buybox market share as in Figure 3b persists.

To disentangle the various factors influencing recommendation status, we estimate a nested logit choice model and report several descriptive facts.[12] To

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.

[11] We obtain this additional data from Keepa, a price tracker for goods on Amazon.

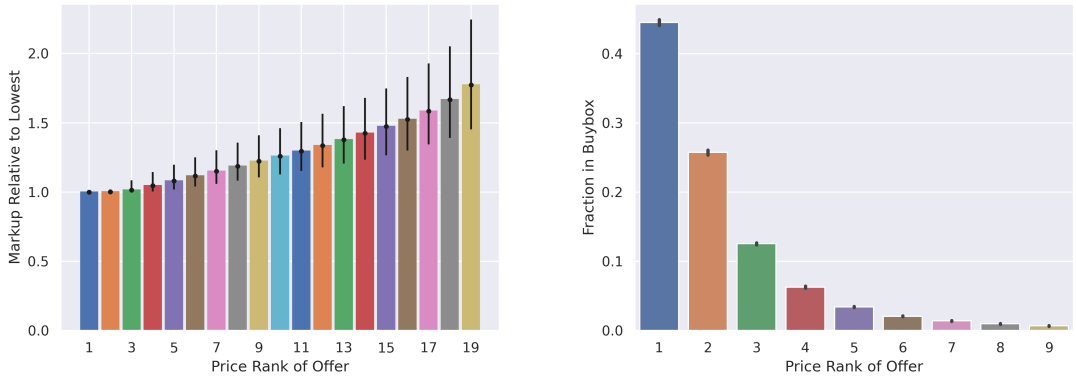
[12] We relegate the technical discussion of this model to Section 5 below, including our handling of potential price endogeneity. For now, we seek to describe, not to ascribe causality.

facilitate exposition, we transform coefficients to answer questions of the following form: if a seller were to move from the 1st percentile of feedback count to the 99th, how much more could she charge while keeping the same probability of being recommended as before? In Table 2, we show that she could raise her price by 1.28%. The effect of increasing positive feedback is about the same, but dispatch time matters more: after moving from the 1st to the 99th percentile of the dispatch time distribution, a merchant would have to *decrease* her price by 9.91% to maintain her old recommendation share. In addition, there are benefits to utilizing the platform's fulfillment network, above and beyond the benefits to dispatch time: switching to FBA allows merchants to raise their prices by 12.36% without losing recommendation share. This benefit naturally also accrues to offers made by the platform itself. However, these offers benefit by an *additional* 5.70%, i.e., a total of 18.06%. While this additional benefit may inspire accusations of self-dealing, there are legitimate reasons to emphasize the platform's own offers if the customers prefer these offers. Such a preference may emerge if consumers are worried about counterfeit goods, or prefer how the platform handles 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, though 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, the effect of price depends on the MSRP of the product: for product-pages with higher MSRP, the platform is more willing to recommend higher-priced offers. Indeed, the recommendation effect is approximately linear in price divided by MSRP, supporting the functional form assumption we make below.

4 Model

Our descriptives reveal determinants of an offer's recommendation status, and hint at consumer preferences. However, more structure is needed to examine how alternative recommender systems affect demand, entry, pricing, and market power. To this end, 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



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

Figure 3: Prices & Recommendation Status Vary Substantially 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. All ties are broken randomly.

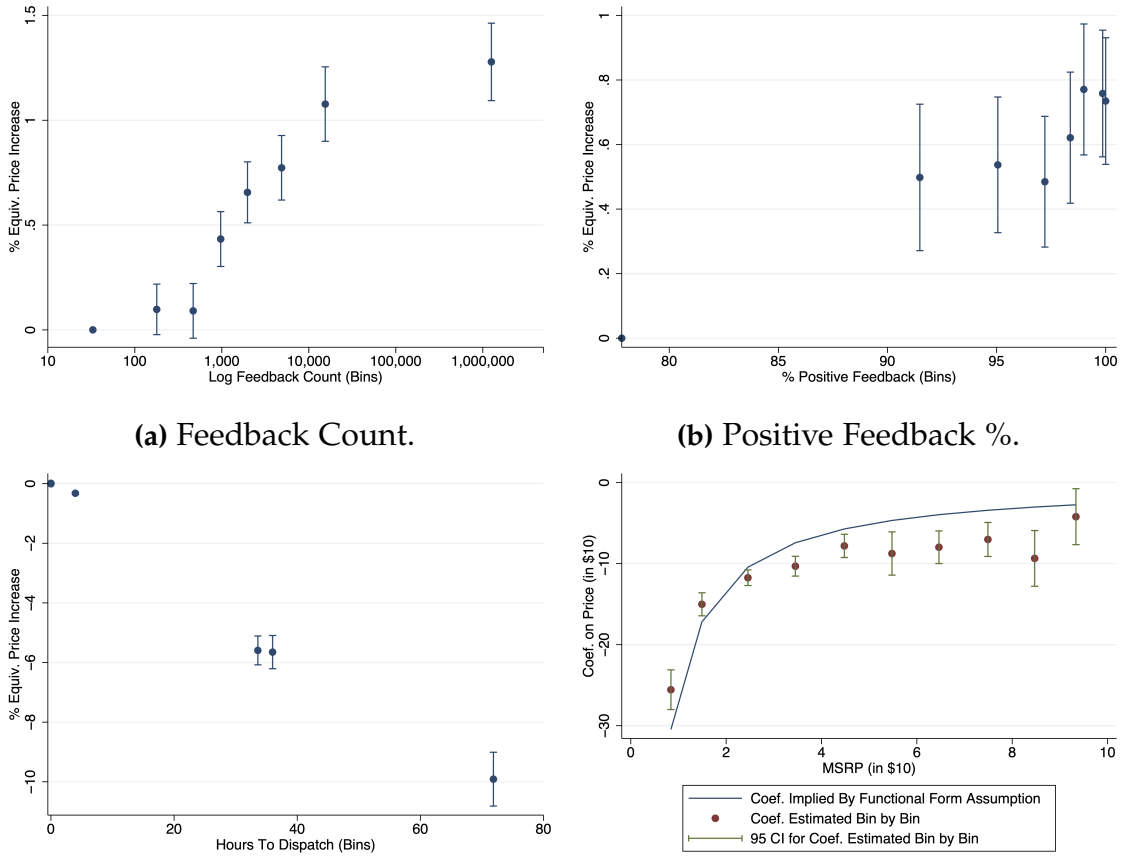
approach focuses on the platform's ability to guide search by influencing customers' consideration sets (Goeree 2008). While past literature has employed a similar approach to investigate platform-guided search on eBay (Dinerstein et al. 2018), our key modeling contribution is our entry model. By modeling entry, we let the data speak to the platform's power by explicitly accounting for its need to attract agents on both sides of the market.

This section proceeds as follows. First, we specify demand: Consumers have nested logit preferences with endogenous consideration sets. Next, we consider how the platform's recommendation algorithm, also a nested logit, maps offer characteristics to a recommendation. Finally, we specify supply: on each product, knowing their marginal costs, sellers enter, observe their opponents, then play a Bertrand-Nash pricing game.

4.1 Consumer choice

Fix a market t, which is a product on the platform marketplace. Where obvious, we will suppress the market subscript for ease of notation.

Within each market, demand follows a standard (nested) logit framework except



(a) Feedback Count.

(b) Positive Feedback %.

(c) dispatch Time.

(d) Price Coefficient.

Figure 4: Non-Price Characteristics Give Recommendation (Dis-)Advantage.
Notes: Panels (a)-(c) answer the following question: suppose a merchant were to move from the baseline value of a given feature to the value given on the x-axis; by how much could she raise her price (relative to MSRP) without negatively affecting her recommendation share? Both feedback count and the percentage of positive feedback matter slightly, but time until dispatch (measured here in hours) matters a lot more (note the scale of the y-axis.) Panel (d) (note different y-axis) illustrates how the price coefficient for a product-page varies with the MSRP of said product, justifying our later use of price divided by MSRP instead of raw price as an explanatory variable. The solid line in panel (d) illustrates what a constant coefficient on Price/MSRP implies in terms of price coefficients across the MSRP bins. All results are computed using an underlying nested logit model of recommendation choice discussed in Section 5. However, the model estimated for these graphs differs from that of our mainline specification below. Here, we allow for non-linear effects of the covariates by including dummies for quantile-spaced bins. All standard errors are clustered at the product level.

	Fraction of Notifications		
	Overall	FBA	Amazon
FBA Offer Exists	97.31%	100.00%	100.00%
Amazon Offer Exists	18.41%	18.92%	100.00%
Recommended Offer Is ...			
Lowest Priced	41.63%	41.32%	45.02%
Second Lowest Priced	26.98%	27.23%	30.84%
Highest Feedback Count	29.24%	29.39%	64.73%
Highest Feedback Rating	20.25%	20.22%	9.89%
Fastest Dispatch	26.43%	26.48%	31.70%
Lowest Priced FBA	51.40%	52.83%	64.59%
FBA	86.52%	88.92%	97.12%
Amazon	11.91%	12.24%	64.73%
Lowest Fastest Dispatch	52.95%	53.36%	65.18%

Table 1: Determinants of Recommendation Status.

Notes: Each entry in this table gives the fraction of recommendation observations satisfying the criteria listed in the first column for various subsets of the data. The second column ("Overall") is based on the entire set of observations. The third column ("FBA") employs only observations for which there is an offer that is fulfilled by Amazon and the fourth column ("Amazon") uses only observations for which there is an offer by Amazon itself. All ties are broken randomly.

that consumers form endogenous consideration sets as in Goeree (2008). There are two types of consumers. A fraction ρ are "sophisticated." Sophisticated consumers ignore the recommendation 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 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 ξ_{jt}:

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

	Mean	Std. Dev.	1st	99th	Rec. Effect	R^2 on Offer FE
Is FBA?	0.61	0.49	0	1	12.36% (0.47)	0.98
Is Amazon?	0.02	0.14	0	1	5.70% (0.25)	1.00
Feedback Count	160,102.72	938,659.14	2	6,506,262	1.28% (0.09)	1.00
Pos. Feedback %	94.86	10.32	62	100	0.73% (0.10)	0.90
Dispatch Time	18.15	25.76	0	84	-9.91% (0.46)	0.97

Table 2: Observable Quality Covariates & Their Effect on Recommendations.
Notes: We provide the distribution of observable quality covariates (columns 2-5). Column 6 reports the amount by which a merchant would have to increase price (as % relative to MSRP) to keep her recommendation share constant, after moving from the 1st percentile to the 99th percentile of each covariate. The last column provides the R^2 from regressing these covariates on offer fixed-effects, indicating the fraction of variation that can be attributed to the cross-section. Standard errors (for column 6) are clustered at the product level.

This formulation allows the price coefficient α_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: $\delta_{0t} = 0$.

An important feature of our setting is that the inside options available to consumers are very similar, and hence plausibly quite substitutable: they all correspond to getting the same product, albeit delivered via different channels at varying speed from sellers of varying reputation. Hence, plausibly, the inside alternatives are more substitutable with each other than with the outside option, which includes both buying other products on the platform and buying the product on some other platform or offline. The outside option also includes not buying the product at all.

To reflect this difference in substitutability, we partition the set of products as $\{0\} \cup \mathcal{J}_t$ and let $g = 1$ be the index of the nest of inside options. The utility

consumer i derives from product j in market t is then

$$v_{ijt} = \delta_{jt} + \zeta_{ig(j)t} + (1-\lambda)\epsilon_{ijt},$$

where ϵ_{ijt} is distributed i.i.d. Type-1 Extreme Valued, and $\zeta_{ig(j)t}$ is common to all options in the same nest. By Cardell (1997), the distribution of $\zeta_{ig(j)t}$ can be specified such that $\zeta_{ig(j)t} + (1-\lambda)\epsilon_{ijt}$ has a Generalized Extreme Value distribution.

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 $j \neq 0$ in market t is

$$\begin{aligned} d_{jt}(j^r) &= \rho \times \frac{\left[\sum_{k\in\mathcal{J}_t} \exp(\delta_{kt}/\lambda)\right]^\lambda}{1+\left[\sum_{k\in\mathcal{J}_t} \exp(\delta_{kt}/\lambda)\right]^\lambda} \times \frac{\exp(\delta_{jt}/\lambda)}{\sum_{k\in\mathcal{J}_t}\exp(\delta_{kt}/\lambda)} \\ &\quad + (1-\rho) \times \mathbf{1}\{j = j_t^r\} \times \frac{\exp(\delta_{jt})}{1+\exp(\delta_{jt})}. \end{aligned}$$

4.2 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.[13] The mean utility δ_{jt}^r of each inside alternative includes observable characteristics $\mathbf{x}_{jt}$, price p_{jt} and the econometrician-unobservable quality ξ_{jt}^r:

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

As with demand, we allow the price coefficient α_t^r to depend on market characteristics and normalize the utility of the outside option, $\delta_0^r = 0$.

Whether due to deliberate randomization or mistakes in evaluation, we will assume that recommendation decisions are based on a shocked version v_{jt}^r of

[13] Formulating the platform recommendation problem in this manner makes it compatible with our discrete-choice demand model. 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, in typical discrete choice demand models, there are usually multiple consumers, each making their own choice in each market.

this mean utility.[14] In particular, shocks follow a nested logit structure like with consumer choice. As the shock structure is perfectly analogous to consumer choice, we will not repeat the details here. Instead, we skip ahead to the implied probability of recommendation:

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

Combining this with our previous result on demand, the market share of offer j is

$$\begin{aligned} s_{jt} \;=\; & \rho \times \frac{\left[\sum_{k\in\mathcal{J}_t} \exp(\delta_{kt}/\lambda)\right]^\lambda}{1+\left[\sum_{k\in\mathcal{J}_t} \exp(\delta_{kt}/\lambda)\right]^\lambda} \times \frac{\exp(\delta_{jt}/\lambda)}{\sum_{k\in\mathcal{J}_t} \exp(\delta_{kt}/\lambda)} \\ & +(1-\rho) \times \frac{\left[\sum_{k\in\mathcal{J}_t} \exp(\delta^r_{kt}/\lambda)\right]^\lambda}{1+\left[\sum_{k\in\mathcal{J}_t} \exp(\delta^r_{kt}/\lambda)\right]^\lambda} \times \frac{\exp(\delta^r_{jt}/\lambda)}{\sum_{k\in\mathcal{J}_t} \exp(\delta^r_{kt}/\lambda)} \times \frac{\exp(\delta_{jt})}{1+\exp(\delta_{jt})}. \end{aligned}$$

4.3 Firm Choices

To close the model, we specify entry and price competition across a set of markets $\mathcal{T}$. Each market $t \in \mathcal{T}$ is associated with a fixed cost F_t and a set of potential entrants $j \in \mathcal{N}_t$. Each putative seller j draws a type $\omega_j \sim G(\cdot)$.[15] She then chooses a pair (χ_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 ϕ of the revenue they generate (as in our application), seller j's payoffs are given by

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

where s_{jt} maps the vectors of types and prices into demand; and A_t is the size of market t, measured in arrivals per month. Our solution concept is Bayes-Nash Equilibrium. We proceed by stating several assumptions for tractability.

[14] We provide evidence for randomization on the Amazon platform in Appendix B.2.

[15] A seller's type ω collects her marginal cost draw, attractiveness to the recommender system (including ξ^r), and attractiveness to consumers (including ξ).

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 (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 5 (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 cross-market cannibalization's role 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 are only a secondary concern.

Assuming unit production costs is common in the literature and particularly plausible in our context as sellers are retailers (not producers). Full information pricing is standard. These assumptions imply that successful entrants play a Bertrand-Nash pricing game (Anderson, De Palma, and Thisse 1992). In our simulations, pricing equilibria always exist but are not necessarily unique. Intuitively, firms will attempt to differentiate by pursuing *either* sophisticated *or* unsophisticated consumers, and different permutations of these roles can create multiple equilibria. Sellers internalize the effect of recommendations when setting prices:

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

Thus, through influencing $\frac{\partial r_j}{\partial p_j}$, the recommendation system partially determines the price elasticity that sellers face — exactly where the platform's power to guide consumer search manifests in our model.

Where there are such multiplicities, we select an equilibrium. We initialize our search for a fixed point in prices by starting the seller with the highest recommender system attractiveness at a lower price than her competitors.[16]

Given a selection rule for the pricing equilibrium, the last two assumptions ensure 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 identities, 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, each firm's only information 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, we could allow entry decisions to depend on private information about vertical characteristics. However, we require the assumption to keep the computation tractable.[17] 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 selection on quality matters, our analysis below may overstate the welfare loss due to recommender-inspired exit. However, given the magnitude of our estimates of quality sensitivity (relative to price

[16] See Appendix C.1 for details. To assess the robustness of this assumption, we also estimated stage games and selected equilibria using alternative rules. These are: (1) a random seller starts at a lower price than others; and (2) the seller with the highest consumer demand attractiveness starts at a lower price. In either case, our results remain qualitatively unchanged.

[17] The entry game is solved many times in an inner loop when we search for parameters in our estimation procedure.

sensitivity), we believe that this is a second-order concern.

Each firm will enter if and only if its expected profits from entering exceed the fixed cost F. Given our assumptions, this expectation depends only on a firm's own unit cost draw. As 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 5 further restricts attention to symmetric equilibria. Thus, all (ex-ante identical) firms share the same (ex-ante) cutoff c^*. In an appendix to a previous version of the paper, conditional on a selection rule for pricing equilibria, we show[18]

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, letting $q_j = x_j'\beta + \xi_j$, 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 π 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 C.2.

5 Estimation

We estimate our model on data covering 50,486 products over a period from 08/26/2018 to 03/25/2020. As we are only informed about which offer is recommended at certain moments (directly following any change in the number of competitors or any competitor's price), and as recommendations potentially vary at a high-frequency, we use only the first fifteen minutes of data after such a notification in the estimation to avoid attenuation bias. We discuss why fifteen minutes in Appendix B.3; the key intuition is that recommendations may vary at high frequency. For the purposes of demand estimation[19], we hence define a market t to be a "product page"-15-minute interval pair, (p, τ).

[18] We omit this proof from this version as the result is very intuitive.

[19] When considering pricing and entry, a market will be a "product page"-month.

In each market, the alternatives $\mathcal{J}_t$ are the various offers on the same product page. This market definition allows us to sidestep estimation of potentially complex preferences over product characteristics: such characteristics do not vary across offers for the same product, so they cancel out in the discrete choice problem. Nevertheless, the offers may still differ in characteristics such as dispatch 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 alternatives offered to customers vary product page by product page. Naturally, preferences over offer characteristics might depend on what the product is. In theory, we could allow preferences to freely vary by product page. In practice, this approach is underpowered. While we observe 64,809 sales, these are distributed across 50,486 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, a difference of \$0.50 may sway a consumer's choice from one offer to another. Yet the relevant price difference for the laptop is typically closer to \$50, not \$0.50. Formally speaking, if R_p is the MSRP of product p, we assume both consumers and the recommendation algorithm evaluate offers based on 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 Figure 4d, this functional form assumption fits our data well. Furthermore, it is supported by our conversations with industry experts.

5.1 The Recommendation Algorithm

We estimate the parameters of the recommendation algorithm by maximum likelihood. Our estimation strategy substantially diverges from prior literature for two reasons. Firstly, we have access to high-frequency, highly 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 2023; Quan and Williams 2018). Secondly, our data comes from a dynamic marketplace characterized by frequent price changes and high merchant turnover. These features suggest that 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[20], 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 these issues, we estimate our parameters via maximum likelihood. Here, we 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, 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, ξ_j^r refers to the unobserved *offer* quality. One may argue that *unobserved* offer quality is unlikely to matter much because we observe all key determinants of offer quality. For instance, we see an offer's time to dispatch, whether Amazon's own logistics operation fulfills the offer, whether the offer is listed by Amazon itself, as well as other measures of seller quality.

This discussion motivates the following identifying assumption:

Assumption 6. *Offer qualities are not correlated with price, i.e.,* $(\mathbb{E}[\xi_{jt}^r|p_{jt}], \mathbb{E}[\xi_{jt}|p_{jt}]) = (\mathbb{E}[\xi_{jt}^r], \mathbb{E}[\xi_{jt}])$.

This assumption seems strong, given the usual intuition in differentiated products demand that higher prices could signal higher unobserved quality. Yet, as there are few dimensions of quality which consumers observe but we do not, we believe this assumption to be plausible. However, while convincing stories that violate Assumption 6 are harder to find in our setting, they exist. For instance, the assumption is violated if consumers use the seller's name to draw inferences about seller quality (e.g., based on prior experiences with the seller) that are not visible to the econometrician.

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

Our data on observable offer characteristics motivate a natural strategy to deal with this potential endogeneity. In Table 2, 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. Hence, we also build estimators that rely on the following alternative identifying assumption:

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

This assumption addresses the potential endogeneity of price by employing 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 B.1. 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 the price coefficient α^r.

We exhibit our estimation results in Table 3. While the first column assumes a (non-nested) logit discrete choice model, the next two specifications employ a nested logit model. Finally, the last column uses the weaker identifying Assumption 7 which is robust to arbitrary correlation of an offer's average unobserved quality with its price.

We begin with a discussion of the effect of our various assumptions on the estimated coefficient on Price/MSRP. Firstly, the most important factor affecting our estimate is correlation in the value of the inside goods: moving from a logit to a nested logit increases the effective inside price coefficient by a factor of two. Including all observable quality measures by moving from (2) to (3) further increases our estimate. However, once we condition on observed quality, switching out our identifying assumption for one robust to arbitrary correlation between average unobserved quality and price (i.e., moving from (3) to (4)) does not seem to affect our estimate much: the implied inside price coefficient is, if anything, lower in (4).

We emphasize a few additional findings.[21] Firstly, the implied price elasticity of the recommendation algorithm is very high, at about -20. This finding is corrob-

[21] Note that our results here vary slightly from our discussion in the descriptive analysis section as the effect of quality is modelled less flexibly here.

orated by our discussions with sellers on the platform. Even more intriguingly, on its seller interface, Amazon explicitly tells the seller which of her "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 $\hat{\lambda} \approx 0.09 < 1$, we find strong evidence that Amazon considers offers more substitutable with each other than with the outside option. This finding is expected as choices of the outside option do not yield any intermediation fees.

Finally, not only does price matter for recommendations; quality does too. Slow shippers are penalized: taking one additional day to dispatch a product is equivalent to having an offer that is 4.6% more expensive. Similarly, FBA offers have a 10.55% price advantage, and Amazon's offers appear to have an (additional) advantage equivalent to a 4.6% price discount. Does this advantage reflect a quality difference, or is it merely self-dealing? Below, our counterfactual results in Section 6.1 support the quality hypothesis.

5.2 Consumer Choice

Our demand estimation face the same challenges of endogeneity and high-frequency data. However, it becomes even more critical not to smooth over short-run temporal variation because of how the fraction of unsophisticated consumers, ρ, is identified. Intuitively, ρ is pinned down by the observed covariance between sales and recommendation status. As this status varies at high frequency, we would smooth over this variation if we were to aggregate over time. Indeed, Appendix B.3 provides evidence that aggregating to time periods longer than 15 minutes attenuates the estimate of ρ. Hence, as before, we proceed by MLE and employ a fixed-effects strategy to address endogeneity.

We face a further 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 use our previous strategy to condition out offer-level fixed effects: when one only observes a single merchant's sales, it is impossible

	(1)	(2)	(3)	(4)
Price / MSRP	-7.54 (0.093)	-2.05 (0.073)	-2.18 (0.070)	-23.55 (1.743)
Dispatch Time (in Days)			-0.10 (0.005)	
100 $\times$ Log(#Feedback+1)			0.39 (0.024)	
Fulfilled by Amazon?			0.23 (0.010)	
Sold by Amazon?			0.10 (0.005)	
Constant (Inside Options)	9.64 (0.121)	4.48 (0.075)	4.48 (0.069)	
Nesting Coefficient (λ)		0.14 (0.005)	0.09 (0.003)	
Offer FE?	✗	✗	✗	✓
Elasticity	-6.62	-12.78	-20.40	
Inside Price / MSRP	-7.54	-14.64	-24.22	-23.55

Table 3: Recommender System.

Notes: Estimates from maximum-likelihood estimation of the recommender system using data on 50,486 product pages spanning 49,069 unique merchants between 2018-08-26 01:18:08 and 2020-03-25 23:39:41. Each product page was observed for an average of 390.32 15-minute periods. The reported coefficients measure the effect of price (normalized by the MSRP), the effect of an extra day of time until dispatch, the effect of a one percent increase in number of feedback, the effect of an offer being fulfilled by Amazon, and the mean utility of the inside option (relative to the outside option). λ is the nesting coefficient (with $\lambda = 1$ corresponding to no nesting). "Elasticity" refers to the average price elasticity of the recommender system, and "Inside Price/MSRP" refers to the implied coefficient on Price/MSRP in the choice between inside options (calculated by dividing the Price / MSRP coefficient by the nesting coefficient.) Our preferred model is (3), which accounts for observable quality covariates (and hence price endogeneity); models (1) and (2) are exhibited for illustration and (4) to show that controlling for quality more flexibly using offer fixed-effects does not impact estimates of the (inside) price coefficient. All standard errors are clustered at the product level.

to condition on pairs of periods such that a sale goes to one firm on the first and another firm on the second.

In Appendix B.1, we show how we can exploit the high-frequency nature of our data to condition out fixed-effects nonetheless. Still, the alternative estimator derived there (unlike the previous section) lacks a clean interpretation when demand has a nesting structure. In particular, we focus on pairs of periods between which all except the observed merchant's offer characteristics (including prices) stay fixed. Then, we compare the sales our observed merchant makes in periods where her price is high to her sales in periods where her price is low. However, in doing so, we mostly capture substitution between our merchant and the outside option (as sales are a rare event).

To estimate the market size, we employ additional data on sales ranks. Since their introduction into the literature in Chevalier and Mayzlin (2006), sales ranks on Amazon have frequently proxied for total product-level sales (e.g., Reimers and Waldfogel 2019, De los Santos, O'Brien, and Wildenbeest 2021). This practice has diffused from academia to industry, with companies offering ranks-based sales estimates as a service to merchants. For our mainline results, we utilize data from one of these companies: we obtain the total estimated sales on a given product page by plugging sales ranks into the AMZScout Sales Estimator[22]. Finally, we translate this figure to a market size by dividing by the percentage of sessions on Amazon that result in sales: 12.3% (digitalcommerce360.com). In Appendix B.4, we perform two robustness checks[23] for this procedure.

Before we discuss results, a final identification concern involves the potential endogeneity of recommendations. As emphasized by Ursu (2018), recommender systems typically recommend offers likely to sell. Thus, if Amazon knew an offer's unobserved quality ξ_{jt}, the company could be expected to recommend offers with high ξ_{jt} more frequently. Formally, this would imply $\mathbb{E}[\xi^r_{jt}|\xi_{jt}]$ is increasing in

[22]To be exact, we collapse our data to the product-month level by taking the mean sales rank. For each product-month, we estimate total sales by interpolating on a dataset that provides the AMZScout estimates on a separate log-spaced grid for each top-level product category; we are grateful to Gutiérrez (2021) for sharing these data. Finally, we average across months to form a final estimate of a product's total sales.

[23]Firstly, we verify that the AMZScout data agrees with another industry source, JungleScout; the estimated relationships are not identical but very similar. Secondly, we employ our own data on sales to estimate a linear relationship between observed log sales and log sales ranks. The resulting demand estimates utilizing this alternative measure of market size are qualitatively indistinguishable from those using the industry sources.

ξ_{jt}. If we do not account for this endogeneity, we could overstate the impact of recommendations. Indeed, our estimation as discussed so far is based on

Assumption 8. *Unobserved offer quality for consumers is independent of unobserved recommendation offer quality, i.e.,* $\mathbb{E}[\xi^r_{jt}|\xi_{jt}] = \mathbb{E}[\xi^r_{jt}]$.

However, this assumption is testable. In particular, we exploit Buybox rotations to make progress. Even if offer characteristics remained constant over time, recommendations (unevenly) rotate frequently between sellers. This rotation creates situations in which an offer that is very unlikely to be recommended receives a lucky draw (of ϵ^r_{jt} in our model) and is recommended for a short period. Intuitively, as long as offer quality is invariant over time (Assumption 7), these surprise recommendations identify the causal effect of being recommended. In Appendix B.2, we verify that our model — estimated under Assumption 8 — successfully matches the increase in sales from recommendation surprises. This confirms Assumption 8 under the maintained Assumption 7 (i.e., that quality is time-invariant.)

Our consumer choice estimation results can be found in Table 4. Our preferred specification (7) incorporates our consideration set model (i.e., does not restrict ρ) and observable quality covariates. The overall price elasticity of demand is around -11. By comparing the price elasticity holding recommendations fixed to the one where recommendations are allowed to vary, we conclude that the recommendation algorithm intensifies price competition by increasing the elasticity by 39.87%.

Regarding non-price characteristics, more feedback appears not to affect demand much. However, consumers prefer faster dispatch, valuing a day shorter time to dispatch about the same as lowering price by 6.90%. Similarly, an FBA offer is valued the same as a 8.62% decrease in price, and Amazon offers are valued as being worth a 29.31% discount. We caution, however, that this effect is estimated exclusively from demand data on offers not listed by Amazon: the demand for Amazon offers is inferred by observing the "shadow" that Amazon casts in terms of reduced sales on our third-party merchants.

We now discuss the alternate specifications exhibited in columns (1) to (6) of Table 4. 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., they assume all consumers are sophisticated. By

comparing models (1) to (2) and (3) to (4), we conclude that the inclusion of observed quality covariates only slightly moves the estimated price coefficient. This finding suggests that there is no or little endogeneity of price with respect to these observed covariates, suggesting that the correlation of unobserved quality with price is also likely to be small.

To assess this hypothesis, we compare a specification with offer-level fixed effects (5) to those without (1–4). Amongst the non-nested models, we find that (1) and (2) yield higher estimates of price sensitivity than (5). This result goes opposite to what we would expect with the usual endogeneity concern (where price is positively correlated with quality). However, as suggested above, our fixed-effects estimator captures mostly substitution between the inside and outside good. Hence, we prefer to compare (5) to columns (3) and (4). In these models, the price coefficient captures the extent of substitution between the inside and outside option (dividing said coefficient by the estimated λ would yield the coefficient that governs substitution between the inside options.) Our estimate in (5) is between those of (3) and (4) – a result we would expect if price endogeneity was not a major concern.

The second set of columns (6–7) drops the restriction that $\rho = 1$. When ρ is freely estimated, all our models agree that $\rho < 1$, i.e., a substantial fraction of unsophisticated consumers only considers the recommended option. We find that $1 - \rho \approx 34\%$ of consumers are "naive": 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 who complete Amazon purchases in three minutes or less: 28% (aboutamazon.com). By contrast, we caution against comparing ρ to a common figure that about 80% of sales go 'through' the Buybox. This statistic is purely correlational. Crucially, a sophisticated consumer may frequently find herself purchasing the offer that Amazon recommends to her, not because it is recommended, but because it is a good offer. In fact, this purchase is almost guaranteed because both the recommendation algorithm and the consumer are very price sensitive. Furthermore, consider estimating what fraction of sales go to offers that are in the Buybox at the time of the sale. This estimate will be sensitive to the number of offers on a given product. At the extreme, when a product has only one offer, as long as the recommendation is not canceled (a rare occurrence), by definition 100% of sales go through the

Buybox. By contrast, in our data, just under 50% of sales are sales of offers that are in the Buybox at the time of the sale. Since our data are sourced from a repricing company, we disproportionately see products with multiple competitors, which would lead to the discrepancy between the industry figures and ours. Note, however, that this oversampling of products with multiple merchants does not bias our estimates of ρ as long as our model is correctly specified.

5.3 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 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, our main technical contribution to the literature is to render it feasible.

For illustration, fix a market t. To ensure there can be no negative realizations of wholesale costs, we assume they follow a lognormal distribution, i.e.,

Assumption 9 (Unit Costs DGP). $c_{jt} = (\theta_0^c + x_t'\theta_x^c)\epsilon_{jt}, \quad \epsilon_{jt} \sim LogN(0, \theta_\sigma^c)$.

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 product's MSRP in x_t.

Similarly, we assume fixed costs follow a mixture of lognormal distributions:

Assumption 10 (Fixed Costs DGP). $F_t \sim LogN(\theta_0^F + x_t'\theta_x^F, \theta_\sigma^F)$ *with probability* ω*, and* $LogN(\underline{\theta}_0^F, \underline{\theta}_\sigma^F)$ *otherwise.*

Here, $\theta^F = (\theta_0^F, \theta_x^F, \theta_\sigma^F, \omega, \underline{\theta}_0^F, \underline{\theta}_\sigma^F)$ are parameters that can be estimated. In our data, the distribution of the number of entrants per market appears bimodal: some markets have very few entrants while others have fifteen or more. This pattern motivates our bimodal specification for fixed costs. Some products, like Dickies socks, are easier to source; others, like LG refrigerators, are less so. Since estimating the distribution of fixed costs is less computationally costly, we allow the mean of the high fixed cost component to vary with the suggested retail price and the market size.

	No Rec. Sys.					Rec. Sys.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Price / MSRP	-1.37 (0.152)	-1.42 (0.161)	-0.61 (0.161)	-0.68 (0.136)	-0.63 (0.182)	-0.51 (0.155)	-0.58 (0.134)
Dispatch Time (in Days)		-0.54 (0.096)		-0.05 (0.013)			-0.04 (0.011)
100×Log(#Feedback+1)		2.53 (2.068)		-0.79 (0.187)			-0.82 (0.220)
Fulfilled by Amazon?		0.24 (0.177)		0.05 (0.016)			0.05 (0.017)
Sold by Amazon?		4.75 (0.267)		0.16 (0.037)			0.17 (0.045)
Constant (Inside Options)	-3.49 (0.155)	-3.60 (0.321)	-2.60 (0.126)	-2.53 (0.140)		-2.62 (0.132)	-2.53 (0.142)
Nesting Coefficient (λ)			0.06 (0.020)	0.06 (0.012)		0.06 (0.021)	0.05 (0.014)
Sophisticates Fraction (ρ)						0.60 (0.018)	0.66 (0.016)
Offer FE?	✗	✗	✗	✗	✓	✗	✗
Elasticity (Rec. Fixed)	-1.44	-1.50	-8.21	-10.07		-6.81	-8.36
Elasticity (Rec. Vary)						-9.04	-11.33

Table 4: Demand Estimation Results.

Notes: Estimates from maximum-likelihood estimation of demand using data on 50,486 product pages spanning 49,069 unique merchants between 2018-08-26 01:18:08 and 2020-03-25 23:39:41. Each product page was observed for an average of 390.32 15-minute periods associated with 64,809 sales. Reported coefficients measure the effect of price (normalized by the MSRP), the effect of an extra day of time until dispatch, the effect of a percentage increase in number of feedback, the effect of an offer being fulfilled by Amazon, and the mean utility of the inside options (relative to outside option). λ is the nesting coefficient for sophisticated consumers and ρ is the fraction of consumers that are sophisticated. Elasticity (Rec. Fixed) refers to the average price elasticity of demand assuming that price does not influence consideration sets. Elasticity (Rec. Vary) refers to the same using the earlier estimates of the recommender system according to which it does. All standard errors clustered at the product level.

Identification is standard. The entrant distribution allows us to draw conclusions about the fixed cost parameters θ^F: the mean number of entrants is informative about θ_0^F; its covariance with market characteristics tells us about θ_x^F; its variance speaks to θ_σ^F. Finally, the number of products on which there are high numbers of entrants speaks to the mixing coefficient ω.[24] Similarly, the price distribution contains all the required information about wholesale cost parameters: the mean price across markets provides a moment that speaks to θ_0^c; its covariance with the suggested retail price gives us θ_x^c; and the within-market price variance tells us about θ_σ^c.

Estimation is computationally challenging. To evaluate the moments for given parameters θ, we must solve the entry game for each of 50,486 markets. However, to solve a single entry game, we must find its associated wholesale cost cutoff c_t^*. Doing so requires us to compute equilibrium profits under many candidate wholesale cost values. Finally, we must evaluate many candidate parameters θ during our outer GMM optimization procedure.

We employ several techniques 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 ζ-markup equation of Morrow and Skerlos (2011, henceforth MS), which has stronger local convergence properties[25] (Conlon and Gortmaker 2020). Moving up to the entry game, 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[26] optimization algorithm (Cartis et al. 2019, henceforth DFO-LS). This solver separately models the response of each moment to each parameter, delivering performance superior to the classic simplex algorithm (Nelder and Mead 1965). Nevertheless, a full run of our estimation procedure takes about 64 hours on a 24-core machine.

[24] In practice, for many values of the mean and variance coefficients of the low fixed cost shock $(\underline{\theta}_0^F, \underline{\theta}_\sigma^F)$, all merchants on said product enter the market. Thus, we calibrate $\underline{\theta}_0^F$ to -5, a very low value; and constrain $\underline{\theta}_\sigma^F = \theta_\sigma^F$.

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

[26] 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.

The results from this estimation procedure are available in Table 5. For successful entrants, we find wholesale costs of approximately 87% of the manufacturer's suggested retail price, which seems plausible for small retailers.

We estimate an interquartile range for fixed costs of [US$0.49, US$1.63]. While our fixed costs appear small, they are consistent with the low barriers to entry on Amazon. For example, 59% of merchants set prices *higher* than MSRP, with 15% pricing at higher than 1.5 times MSRP. Since the recommender system often directs consumers to the cheapest offer, these high-pricing merchants take very little of the buybox (and hence market) share[27]. However, to account for the entry of these merchants on the competitive fringe, our model settles on low fixed costs: any larger fixed cost would make entry unprofitable for merchants with such high marginal costs.

Having said this, given the empirical sales frequency, our estimated fixed costs seem reasonable. In particular, these values are comparable to storage costs on Amazon. We observe one sale per month for the median offer in our sample. Assume (admittedly heroically) that this means the median merchant is storing a single item (per product), and that this item is less than one cubic foot in size. If this item is stored in an Amazon fulfillment center, the merchant is paying storage fees of $0.69 per product-month (junglescout.com).

6 Counterfactuals

With our estimates in hand, we investigate two questions. First, is Amazon "self-preferencing"? To evaluate this effect, we compare factual outcomes to those of a simulation in which we switch off Amazon's weight on itself in its recommender system: $\beta^r_{Amazon} = 0 < \hat{\beta}^r_{Amazon}$. Second, what is the combined effect of the platform's search guidance through the Buybox? To answer this question, we compare realized outcomes to those of a counterfactual in which recommendations are absent. In this alternate scenario, absent recommendations, naive consumers have to consider a random offer.

[27]This phenomenon of surprisingly high prices is common on e-commerce websites: Dinerstein et al. (2018, p.1855) discuss a similar finding in the context of eBay, and eventually settle on allowing their model to infer high marginal costs for these merchants. Similarly, we allow our model to infer a high variance of marginal cost shocks.

Parameter	Value	Description
θ_0^F	4.90 (0.021)	mean of log fix cost
θ_M^F	2.46 (0.048)	.. interacted with log market size
θ_R^F	1.09 (0.010)	.. interacted with MSRP
θ_σ^F	0.07 (0.010)	variance of log fix cost
ω	0.138 (0.001)	mixing parameter for low fix cost shock
θ_0^c	246.6 (9.320)	coef. of marginal cost on const
θ_R^c	0.97 (0.004)	coef. of marginal cost on MSRP
θ_σ^c	0.343 (0.001)	variance of marginal cost

Table 5: Fixed and Marginal Cost Parameters.

Notes: This table displays wholesale and fixed cost parameter estimates obtained from our simulated method of moments procedure. The fixed cost parameter estimates imply an interquartile range for fixed costs of [US$0.49, US$1.63]. The marginal cost estimates imply a mean wholesale cost of $2.60 + 103% of MSRP for potential entrants, but only wholesale costs of approximately 87% of MSRP for *successful* entrants. All standard errors are clustered at the product level.

We distinguish three time horizons in our counterfactuals. In the short run, we disallow sellers from adjusting prices and whether or not to enter. In the medium run, sellers price optimally but may neither enter nor exit. Finally sellers may change both their prices and entry decisions in the long run.

All results below refer to an estimation sample of 50,486 products, representing approximately US$161 million in revenue and 465 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" (aboutamazon.com). 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 (marketplacepulse.com). Thus, one could extrapolate our results to the entirety of Amazon Marketplace by multiplying our numbers by \$300B/\$161M $\approx$ 1,863 to highlight the magnitude of the issues at stake. Where appropriate, we will be explicit when displaying extrapolated findings; the usual caveats apply.

6.1 "Self-preferencing" 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 6, which summarizes our estimates of the causal effect of self-preferencing on the subset of the market we observe. Thus, when the table reports, e.g., a \$331,632 figure for "Δ Consumer Surplus" in the "Short-Run" column, this indicates that total consumer surplus across our 50,486 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 slightly raises consumer welfare in the short run, by about US$7 on average per product. 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.

Moreover, we find a minimal effect of "self-preferencing" on pricing and entry by third-party merchants in the medium and long run. While there is entry displacement, at -0.02 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 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 "switching on" the platform advantage is $0.7 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 cannot access sales data for Amazon offers, we must infer these sales from estimates of market size and sales on other, non-Amazon offers. Although this procedure yields no formal identification issues, future work can improve on our estimates with better sales data.

Secondly, our model does not account for the platform's pricing strategy: we assume that Amazon's prices remain fixed in all counterfactuals. This assumption is necessary as we believe modeling Amazon as maximizing short-run profits from sales would be inaccurate. To compensate, we investigate in Appendix D.1 to what extent the platform would have to raise prices due to its "self-preferencing" to reverse our conclusion that this practice is welfare enhancing. This reversal happens if Amazon raises prices by more than 2.0%.

6.2 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 natural benchmark captures the case in which they are not provided with a recommendation.[28]

[28]We emphasize that this counterfactual implicitly assumes that consumers would not adjust the size of their consideration sets in response to the absence of recommendations. If this scenario seems implausible, we note that the platform could always make its recommendations compulsory; in particular, it could disallow consumers from interacting with unrecommended offers. Other

	Short-Run	Medium-Run	Long-Run
Prices Adjust?	✗	✓	✓
Entry Decisions Adjust?	✗	✗	✓
Δ Consumer Surplus (CS)	$331,632	$463,640	$465,899
Δ Consumer Surplus (Naive)	$331,632	$415,989	$422,729
Δ Consumer Surplus (Soph.)	$0	$47,651	$43,170
Δ Producer Surplus (PS)	-$65,504	-$80,290	-$82,862
Δ Intermediation Fees	-$150,152	-$74,094	-$84,689
Δ Welfare (CS + PS)	$266,128	$383,351	$383,037
Δ 3P Mean # Sales/Month	-0.15	-0.07	-0.02
Δ 3P Mean Price (% MSRP)		-0.02%	-0.12%
Δ 3P Mean Min Price (% MSRP)		-0.02%	-0.02%
Δ 3P Mean Cost (% MSRP)			-0.09%
Δ 3P Mean # Entrants			-0.02

Table 6: 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 β^r_{Amazon} is the recommendation algorithm's coefficient on the dummy that indicates an offer belonging to the platform operator. Formally speaking, on our sample of 50,486 products, 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 β^r_{Amazon}. In the short run, sellers cannot change their prices or entry decisions. In the medium run, sellers may set prices optimally, but cannot change their entry decisions. Finally, the long-run counterfactual allows sellers to change both their prices and their entry decisions. Outcomes that cannot change (e.g., the mean number of entrants in the medium-run counterfactual) are omitted from the Table for clarity. 3P refers to 'Third-Party', i.e., it indicates that the outcome is computed using only non-Amazon offers.

We exhibit our results in Table 7, 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 gain to consumers of $1863 \times \$16$ million $\approx \$30$ 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 34% 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 prices to fall in the medium run by about 70% of MSRP on average, and by 28% of MSRP for the cheapest offer. The algorithm redistributes surplus from the producer to the consumer side of the economy. Consumers enjoy extrapolated marketplace-wide gains of $1863 \times \$45$ million $\approx$ \$84 billion from search guidance.

As expected, redistributing surplus in a two-sided market affects entry. Indeed, the platform's search guidance is responsible for a reduction of 4.4 entrants/market in the long run. Given an average of 10 merchants per market, this reduction indicates a significant dampening effect on entry. Nonetheless, the merchants that do not enter are on the competitive fringe: with more intense price competition, the mean wholesale cost of entering merchants falls by 17% of MSRP. On net, prices on the platform fall by 93% of MSRP on average, but the lowest prices only fall by 25% of MSRP.

It follows that, relative to what an A/B test would show, the value of the recommendation algorithm is much higher when pricing and entry are considered. Indeed, pricing and entry combined raise the (extrapolated) welfare gains from search guidance technology to $1863 \times \$46$ million $\approx \$86$ billion, much higher than

marketplaces such as ride-hailing services implement such a policy. In this context, our counterfactual world could be understood as one with a random mandatory recommendation only for naive consumers.

	Short-Run	Medium-Run	Long-Run
Prices Adjust?	✗	✓	✓
Entry Decisions Adjust?	✗	✗	✓
Δ Consumer Surplus (CS)	$15,631,914	$44,827,255	$49,180,670
Δ Consumer Surplus (Naive)	$15,631,914	$35,624,925	$42,243,326
Δ Consumer Surplus (Soph.)	$0	$9,202,330	$6,937,344
Δ Producer Surplus (PS)	$1,561,133	-$4,654,287	-$3,650,490
Δ Intermediation Fees	$342,331	$1,476,305	$1,581,494
Δ Welfare (CS + PS)	$17,193,047	$40,172,968	$45,530,180
Δ 3P Mean # Sales/Month	0.50	3.42	8.53
Δ 3P Mean Price (% MSRP)		-69.54%	-93.35%
Δ 3P Mean Min Price (% MSRP)		-28.32%	-24.69%
Δ 3P Mean Cost (% MSRP)			-17.32%
Δ 3P Mean # Entrants			-4.44

Table 7: The Value of Recommendations.

Notes: This table provides the difference in various outcomes, on our sample of 50,486 products, 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 the recommendation algorithm places on various offer features when making its recommendation decision. Formally speaking, for various outcomes x, we report $\Delta_x = \hat{x}(\hat{\boldsymbol{\beta}}^r) - \hat{x}(\mathbf{0})$. Setting $\boldsymbol{\beta}^r = \mathbf{0}$ is equivalent to the platform choosing recommendations uniformly at random (we also forbid the platform from recommending the outside option in this counterfactual). In the short run, sellers cannot change their prices or entry decisions. In the medium run, sellers may set prices optimally, but cannot change their entry decisions. Finally, the long-run counterfactual allows sellers to change both their prices and their entry decisions. Outcomes that cannot change (e.g., the mean number of entrants in the medium-run counterfactual) are omitted from the Table for clarity. 3P refers to 'Third-Party', i.e., it indicates that the outcome is computed using only non-Amazon offers.

the predicted static gains of 1863 $\times$ \$17 million $\approx$ \$32 billion.

Given the gains to price-based search guidance, why does the platform not make its recommendation algorithm even more price elastic? In Appendix D.2, we find that compared to a world with more elastic recommendations, the current algorithm performs better in an A/B test (as over-emphasizing price distracts from other determinants of demand.) However, this conclusion flips once algorithm-induced pricing changes are accounted for. Furthermore, while increased price competition from even more elastic recommendations lowers entry, this entry effect does not outweigh the competition effect. Hence, our estimates suggest that the platform could benefit from making its recommendations even more price elastic.

7 Conclusion

Regulators worry that platforms matching buyers and sellers may influence market outcomes by guiding consumer search through algorithmic recommendations. We explore this issue through a model of intermediation power. Our model enables 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 defaults. On the Amazon marketplace, 34% of customers only consider the default option; the recommendation algorithm raises the price elasticity of demand from 8 to 11. Through their choice of search, ranking, and recommendation algorithms, platforms are indeed able to influence market outcomes, guide consumers toward 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 harms consumers. While we find that the examined recommendation algorithm advantages the platform, this "self-preferencing" positively affects 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 entry. 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, higher competitive pressure lowers entry incentives. While concentrated amongst sellers on the competitive fringe, the reduction in entry nonetheless constrains the extent to which Amazon can drive prices down on the platform.

We discuss future directions. On the demand side, richer sales data can speak to *which* consumers are helped or hurt by platform steering. Future work could also help assess the extent to which our attempt to compensate for the shortcomings of our data by, e.g., proxying for market size has affected our conclusions. On the supply side, our analysis can be adapted to speak to innovation and the growth

of platforms over the long-term. In the very long run, merchant quality and investment matter; how these dynamics interact with recommender systems is an open question. Finally, while platforms are large, they still have competitors. While our entry margin implicitly accounts for the ability of merchants to switch platforms, a richer model could 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: *Quantitative Marketing and Economics* 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.

Aguiar, Luis, Joel Waldfogel, and Sarah Waldfogel (2021). "Playlisting Favorites: Measuring Platform Bias in the Music Industry". In: *International Journal of Industrial Organization* 78, p. 102765.

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, James Levinsohn, and Ariel Pakes (1995). "Automobile Prices in Market Equilibrium". In: *Econometrica*, pp. 841–890.

Berry, Steven T (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.

Bodoh-Creed, Aaron L, Joern Boehnke, and Brent Hickman (2021). "How Efficient are Decentralized Auction Platforms?" In: *The Review of Economic Studies* 88.1, pp. 91–125.

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 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.

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.

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.

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

Chen, Nan and Hsin-Tien Tsai (2023). "Steering via Algorithmic Recommendations". In: *RAND Journal of Economics*.

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.

Ciotti, Fabrizio and Leonardo Madio (2023). *Competition for Prominence*. Mimeo.

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.

Crawford, Gregory et al. (2022). "Amazon Entry on Amazon Marketplace". In: *CEPR Press Discussion Paper No. 17531*.

De Los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest (May 2012). "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior". In: *American Economic Review* 102.6, pp. 2955–80.

De los Santos, Babur, Daniel P O'Brien, and Matthijs R Wildenbeest (2021). "Agency Pricing and Bargaining: Evidence from the E-Book Market". In: *Kelley School of Business Research Paper* 18-90.

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 (May 2021). *District of Columbia vs Amazon Inc*.

Donnelly, Robert, Ayush Kanodia, and Ilya Morozov (2023). "Welfare Effects of Personalized Rankings". In: *Marketing Science*.

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.

Etro, Federico (2022). "The Economics of Amazon". In: *Available at SSRN 4307213*.

Farronato, Chiara, Andrey Fradkin, and Alexander MacKay (2023). "Self-Preferencing at Amazon: Evidence from Search Rankings". In: *AEA Papers and Proceedings* 113, pp. 239–43.

Gandhi, Amit, Zhentong Lu, and Xiaoxia Shi (2023). "Estimating Demand for Differentiated Products with Zeroes in Market Share Data". In: *Quantitative Economics* 14.2, pp. 381–418.

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.

Gutiérrez, Germán (2021). *The Welfare Consequences of Regulating Amazon*. Working Paper.

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*, pp. 649–670.

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

Lam, Honn Tai (2021). *Platform Search Design and Market Power*. Working Paper.

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

Matzkin, Rosa L. (2007). "Chapter 73 Nonparametric Odentification". In: ed. by James J. Heckman and Edward E. Leamer. Vol. 6. Handbook of Econometrics. Elsevier, pp. 5307–5368.

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*, 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.

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 (Sept. 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*. Report, 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.

Raval, Devesh (2022). "Steering in One Click: Platform Self-Preferencing in the Amazon Buy Box". In: *Unpublished manuscript*.

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.

Seim, Katja (2006). "An empirical model of firm entry with endogenous product-type choices". In: *The RAND Journal of Economics* 37.3, pp. 619–640.

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.

U.S. Congress (2021). *American Choice and Innovation Online Act*.

U.S. Senate Judiciary Committee (2020). *Investigation of Competition in Digital Markets*. Tech. rep. Congress.

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

Ursu, Raluca M. (2018). "The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions". In: *Marketing Science* 37.4, pp. 530–552.

Online Appendix

A Data Description

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 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., price updates, exits or entries). Secondly, the company is informed of any sales 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 B.4 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 offer, we observe (i) the recommendation status, (ii) whether the offer belongs to a "featured" merchant, (iii) whether Amazon fulfills it, (iv) its list price and (v) shipping price, (vi) the feedback count, (v) the fraction of positive feedback for the associated seller, and (vi) the time to dispatch. Whenever we refer to price, we mean the sum of list price and shipping price. The main challenge of this dataset is its 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. Though sales data linked to sellers on Amazon are rare, a caveat with 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." Being a "featured merchant" (FM) is a hard prerequisite for a merchant to become eligible to be recommended. To become a FM, the seller has to pay $40 a month to Amazon for a professional seller account. Second, for each product, we need a proxy for 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. Third, we subset to notifications such that there are no additional notifications for the next fifteen minutes on the same product page; this restriction allows us to avoid taking a stance on how to aggregate over time by focusing on "stable" markets. Finally, we merge our data by associating each fifteen-minute post-notification period with the sales we observe in that period.[29]

After merging, we obtain a dataset at the 'offer x 15m period'-level with 171,005,258 observations. As our data does not contain reliable information about the feedback count for Amazon, we replace Amazon's feedback count with the maximum value in the dataset.[30] Summary statistics for our final dataset are provided in Table 9(a), and further information (e.g., on the number of unique products) is provided in Table 9(b). For instance, from August 25, 2018 to March 26, 2020, we observe 64,809 sales across 50,486 products which we can attribute to a notification at most fifteen minutes before the sale.

B Recommendation & Consumer Choice Estimation

B.1 Price Endogeneity

Our mainline estimates make use of identifying Assumption 6 which requires $\mathbb{E}[\xi_{jt}|p_{jt}] = \mathbb{E}[\xi_{jt}]$. However, a typical worry in demand estimation is that unobserved quality ξ_{jt} and price p_{jt} may be correlated. Hence, we also develop estimators that rely on quality merely being time-invariant, i.e., Assumption 7. These estimators allow for arbitrary correlation of an offer's unobserved quality and price, as long as quality and price do not co-move within the same offer over time. This assumption is plausible because *observable* offer quality is mostly

[29] This amounts to taking an extreme stance in a bias-variance tradeoff in favor of 'no bias': we observe many more sales (on the order of two million) than we use in estimation, but these sales happen in periods where we would have to guess the recommendation status of the offers on the market. See Appendix B.3 for evidence on how expanding the data used in estimation would yield attenuation bias.

[30] This cannot influence our demand or recommendation predictions, as we are controlling for an Amazon dummy in the relevant models. However, if we are significantly underestimating the feedback count of Amazon, it could influence the interpretation of the Amazon coefficient in, e.g., the recommendation estimation. In such a case, if we were to observe the truth, the estimated coefficient on Amazon would be reduced.

	N	Avg.	Std. Dev.	Min.	50th	Max.
Price/MSRP	171,005,258	1.18	0.42	0.38	1.07	4.71
FBA	171,005,258	0.61	0.49	0.00	1.00	1.00
Amazon	171,005,258	0.02	0.14	0.00	0.00	1.00
Log (Feedback Count + 1)	171,005,258	7.38	2.79	0.00	7.22	15.69
Dispatch Time (in Days)	171,005,258	0.76	1.07	0.00	0.00	3.50
Recommendation Status	171,005,258	0.11	0.31	0.00	0.00	1.00
Sales (15m)	19,705,784	0.00	0.08	0.00	0.00	32.00
Sales (1h)	19,705,784	0.01	0.19	0.00	0.00	63.00
Sales (24h)	19,705,784	0.34	2.52	0.00	0.00	471.00

(a) Offer-Level Summary Statistics.

# Unique Products	50,486
# Merchants	49,069
# Markets (Product x Time)	19,705,784
# Sales (15m)	64,809
# Sales (1h)	270,753
# Sales (24h)	6,784,567
# Markets with Sales (15m)	51,883
# Markets with Sales (1h)	187,348
# Markets with Sales (24h)	2,224,660
Earliest Market	2018-08-26 01:18:08
Latest Market	2020-03-25 23:39:41
Earliest Market w/ Sale	2018-08-26 02:06:39
Latest Market w/ Sale	2020-03-25 20:51:41

(b) Descriptives

Table 9: Summary Statistics and Descriptives.

Notes: These tables displays summary statistics for the dataset used in recommendation and consumer choice estimation. For the left table, the unit of observation is an offer on a given market, so a (merchant, 15m period, product page). As discussed in Appendix B.3, we are informed about the offer characteristics at time t but observe sales whenever they happen. Hence, we experiment with different ways of aggregating the data: 'Sales (15m)' measures the sales of an offer in the 15 minutes after we receive information about offer characteristics, 'Sales (1h)' for the next hour, and 'Sales (24h)' for the next 24h. If multiple notifications arrive within 24h, they could be associated with the same later sale; this problem does not emerge for 'Sales (15m)' (we condition on notifications at least 15m apart). The right table displays counts of products, merchants, markets, and sales; as well as the first and last markets we see, with or without a sale.

time-invariant. Indeed, the last column of Table 2 reports the R^2 from regressing the quality measures (one by one) on offer fixed-effects. This R^2 is never less than 0.90 and typically above 0.97.

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 for the recommender system estimation). 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.

Recall that the recommender system's mean utility from alternative j at time τ is given by[31]

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

[31]We have fixed some product page p and suppress the relevant subscript. Extending the estimators from one to multiple product pages is trivial.

where in contrast to the main text we have imposed Assumption 7, i.e. offer quality $\xi_j^r \equiv \xi_{j\tau}^r$ 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 $v_{j\tau}^r = \delta_{j\tau}^r + \epsilon_{j\tau}^r$. The recommendation is assigned to the offer with the highest utility index. If $y_{j\tau}^r$ is a dummy for whether an alternative is recommended, we have

$$y_{j\tau}^r = \mathbf{1}\{v_{j\tau}^r \geq \max_k v_{k\tau}^r\}.$$

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,

$$\mathbb{P}\left(y_{j\tau}^r = 1 | y_{j\tau}^r + y_{j'\tau}^r = 1\right) = \frac{1}{1 + \exp\left(\delta_{j\tau}^r - \delta_{j'\tau}^r\right)}.$$

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 periods τ, τ' and consider the event $C = \{y_{j\tau}^r + y_{j'\tau}^r = 1, y_{j\tau'}^r + y_{j'\tau'}^r = 1\}$. Then

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

But neither $\delta_{j\tau}^r - \delta_{j\tau'}^r$ nor $\delta_{j'\tau}^r - \delta_{j'\tau'}^r$ contain ξ_j^r. Thus, to consistently estimate β^r and α^r in the presence of ξ_j^r we can maximize the log likelihood function

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

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

Offer FE under Partial Observability

We observe sales for exactly one alternative j for each product page. Thus we cannot exactly follow the previous subsection: 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-I 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 τ,

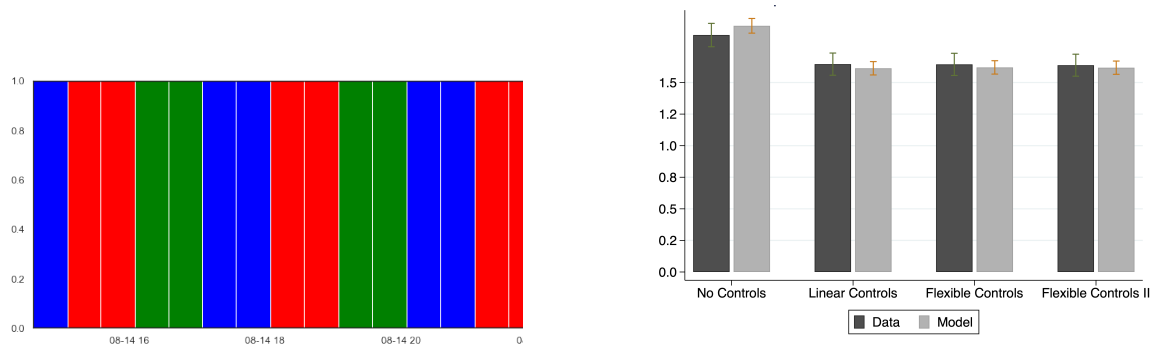
$$y_{ij\tau} = \mathbf{1}\{v_{ij\tau} \geq \max_k v_{ik\tau}\}.$$

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 periods between which all offers for which we do not observe sales remain unchanged. This procedure yields the following likelihood function:

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



(a) The Buybox Rotation. **(b)** Rec. Effect on Demand.

Figure 5: Buybox Rotation and Model Predictions.

Notes: The left panel illustrates Buybox rotations by showing periods in which three distinct sellers of Maxell Lithium Batteries, colored red, green and blue, are recommended. The right panel shows the increase in sales (dark bars) or model-predicted sales (light bars) from being recommended as a fraction of mean sales probability, controlling for offer-fixed effects ("No Controls") or offer-fixed effects and predicted recommendation probability ("Controls"). Our model matches both (i) the correlational relationship between recommendations and sales and (ii) the effect of being recommended 'by surprise', i.e., due to a Buybox rotation.

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.

B.2 Recommendation Endogeneity

As emphasized by Ursu (2018) and Donnelly, Kanodia, and Morozov (2023), a primary concern when estimating the effects of recommender systems on demand is that recommended offers are recommended for a reason. For instance, in our context, offers can be recommended because they are competitively priced — but competitive pricing itself drives sales. Thus, naively comparing sales of recommended offers and non-recommended offers will typically overstate the (causal) power of recommendations: recommended offers sell better not only because they are recommended, but also because they are better along various dimensions (e.g., pricing) valued by consumers.

The preceding discussion suggests a way to disentangle the causal from the correlational effect of being recommended: if we can control for everything consumers value, then any additional impact of being recommended is causal. Implicitly, our mainline estimates above assume that no unobserved factors influence both Buybox status and demand. Therefore, our mainline estimates rely on a "conditioning on observables" strategy for identification.

Additional identifying variation arises from Buybox rotations. In particular, we model the platform's algorithm as solving a discrete-choice problem with extreme-valued shocks. These shocks are ascribed to deliberate randomization. To motivate this modeling choice, we select a product — Maxell Lithium Batteries — sold on the Amazon marketplace. In Figure 5a, we display the identity of the merchant whose offer was recommended by the platform[32] 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, (webretailer.com) 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."

Econometrically speaking, Buybox rotations *exogenously shift* which offer is assigned the Buybox. We can thus compare cases where an offer is assigned to the Buybox for this exogenous reason to cases where an offer was not assigned to the Buybox for this same reason. Formally, a potential threat to identification is that $\mathbb{E}[\xi_{jt}|\xi^r_{jt}] \neq \mathbb{E}[\xi_{jt}]$; in particular, the expectation of ξ_{jt} could be increasing in ξ^r_{jt}. However, suppose we believe Assumption 7 (i.e., that offer quality is time-invariant). Then, since the Buybox rotates over time, these rotations generate exogenous variation in recommendation status. In particular, we can ask: how much more likely is an offer to sell when it is assigned to the Buybox, compared

[32]This figure is made using scraped data from Keepa; while these data have many shortcomings relative to ours, they have the advantage of being sampled independently of price movement. Our data, by contrast, makes it hard to detect rotations as we only see recommendation status when some potential input to recommendations changes.

to when *the same offer* is not assigned to the Buybox and yet has the same model-predicted probability of being assigned to the Buybox?

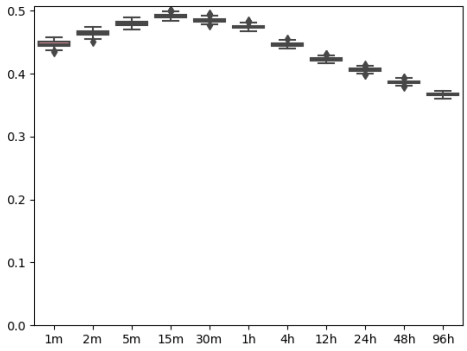
Figure 5b answers this question. The darker bars measure the increase in sales probability from being recommended as a fraction of mean sales probability, through regressions of a sales indicator on recommendation status and controls. In particular, from left to right, we measure this increase using (i) a regression that controls for offer fixed-effects but nothing else; (ii) a regression like (i), but also controlling linearly for the predicted probability of being assigned the Buybox; (iii) a regression like (i), but also nonparametrically controlling[33] for the predicted probability of being assigned the Buybox; and (iv) a regression like (iii), but also controlling for observable offer characteristics directly. Crucially, all of the regressions above include offer fixed effects.

The picture from the dark bars in Figure 5b is clear: even after controlling for offer fixed effects, being recommended raises the probability of a sale. However, the key question is not whether recommendations matter but whether our model correctly picks how much they matter. Hence, we repeat the same regressions with *predicted sale probabilities* as the dependent variable. The results from these regressions are depicted in the lighter grey bars. Since the dark and light grey bars generally overlap, we have suggestive evidence that our model correctly picks up the causal effect of being recommended. Crucially, our model mirrors the reduction in the impact of recommendations on sales after controlling for predicted recommendation status, i.e., when moving from 'No Controls' to 'Linear Controls'. Hence, our model correctly understands both the correlational impact of recommendations and (assuming quality is time-invariant) the causal impact of recommendations.

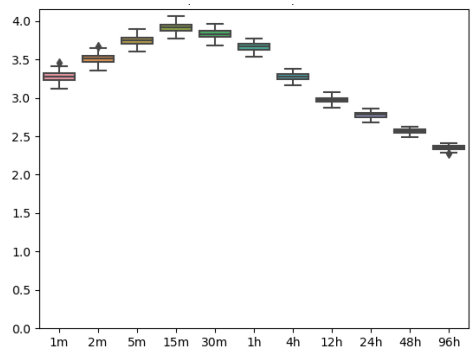
B.3 Aggregation and Attenuation Bias

We only observe the Buybox status of an offer when any offer changes in price or when an offer is added or removed from the set of available offers. This quirk introduces a potential attenuation bias into our estimation: if the Buybox status changes between the last price change and a sale, we may mistakenly conclude

[33] We implement these controls by including separate dummies for 100 quantiles of predicted probability.



(a) Sales Fraction while in Buybox.



(b) Sales Probability Ratio.

Figure 6: Attenuation Bias Reduces Estimates of Buybox Importance.
Notes: The figures depict the fraction of a merchant's sales that happen while she holds the Buybox (left) and the factor by which an offer's sales probability is multiplied if is in the Buybox (right). On the horizontal axis, we vary the time horizon at which we aggregate. For example, the '1h' bin forms estimates based on sales from the first hour after we observe a Buybox change.

that the Buybox does not significantly affect sales. The frequency with which we make these mistakes depends crucially on the frequency with which the Buybox is updated without triggering a notification.

The potential attenuation bias presents us with a bias-variance tradeoff: by concentrating on brief time intervals following notifications, we could reduce or even eliminate the bias generated by using outdated information. However, this approach would substantially limit our effective sample size, as it is improbable for sales to occur within these short periods. Consequently, such a strategy introduces additional variance[34] into our estimates.

As depicted in Figure 6, we find evidence of possible attenuation bias: beyond the initial four points, extending the period after a notification during which we aggregate sales diminishes our estimate of the correlation between sales and Buybox status. The first three points correspond to relatively short periods (1m, 2m, and 15m, respectively) during which *product page caching* may arise[35].

Given Figure 6, we use estimates based on a 15m window after notifications for

[34]This variance is not reflected in Table 10 as it also conditions on periods during which no offer characteristics change, which is less likely at longer time horizons.

[35]Our information on Buybox status reflects the central database, but changes require time to propagate to Amazon product pages, which are often cached for efficiency.

the results in the main text, resolving the bias-variance tradeoff in favor of unbiased estimates. However, in Table 10, we show how our estimates vary as we change the timeframe over which we are aggregating. We find that coefficients vary in the expected way across time horizons. In particular, the longer the time horizon, the larger the estimated fraction ρ of sophisticated consumers (because the correlation between recommendations and sales diminishes as we utilize increasingly outdated recommendation information). However, we also see that the estimates of ρ are already quite stable when comparing the 15m and 1h time horizons. Hence, even if caching did not prevent us from examining even shorter periods, we would expect our results to remain unchanged.

We address one final point: our estimates of the fraction of a merchant's sales that happen while she holds the Buybox are on the lower end compared to industry estimates, which typically range between 70% and 90% (compared to our 51%). This discrepancy may be attributed to our mix of products. Since our data source is a repricing company, products with many merchants are oversampled. Trivially, if a product has only one merchant (and the Buybox is not suppressed), 100% of sales go through the Buybox. Similarly, the more merchants a product has, the fewer sales will go through the Buybox.

B.4 Market Size

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.[36] As we are estimating an entry model, we must 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.[37] To combat this problem, we exploit two strategies, each based on a finding in Chevalier and Mayzlin (2006), built on in De los Santos, O'Brien, and Wildenbeest (2021) and Reimers and Waldfogel (2019): (log) sales ranks are excellent proxies for (log) sales.

Our first strategy utilizes industry data. Merchants frequently need to know

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

[37] This is essentially the problem discussed in Gandhi, Lu, and Shi (2023), only that it arises for us mostly because of high-frequency data and partial observability, and that sales rank data allows us to employ a more tailor-made fix.

	Own Scale			AMZScout Scale		
	(1)	(2)	(3)	(4)	(5)	(6)
Time Horizon	15 Min.	1 Hour	24 Hours	15 Min.	1 Hour	24 Hours
Price / MSRP	-0.77 (0.158)	-0.95 (0.086)	-1.40 (0.105)	-0.58 (0.134)	-0.77 (0.096)	-1.22 (0.127)
Dispatch Time (in Days)	-0.05 (0.014)	-0.07 (0.011)	-0.12 (0.032)	-0.04 (0.011)	-0.06 (0.010)	-0.06 (0.024)
100 $\times$ Log(#Feedback+1)	-0.38 (0.167)	-0.32 (0.174)	0.19 (0.608)	-0.82 (0.220)	-0.97 (0.172)	-1.09 (0.483)
Fulfilled by Amazon?	0.03 (0.017)	0.04 (0.017)	0.18 (0.044)	0.05 (0.017)	0.06 (0.015)	0.22 (0.054)
Sold by Amazon?	0.18 (0.045)	0.20 (0.027)	0.54 (0.157)	0.17 (0.045)	0.20 (0.032)	0.42 (0.120)
Constant (Inside Options)	-0.58 (0.163)	-0.51 (0.118)	-0.54 (0.108)	-2.53 (0.142)	-2.44 (0.116)	-2.36 (0.124)
Nesting Coefficient (λ)	0.07 (0.016)	0.08 (0.008)	0.23 (0.059)	0.05 (0.014)	0.06 (0.009)	0.18 (0.049)
Fraction of Sophisticates (ρ)	0.65 (0.013)	0.67 (0.009)	0.77 (0.015)	0.66 (0.016)	0.69 (0.012)	0.78 (0.016)
Elasticity (Rec. Fixed)	-8.67	-9.33	-4.49	-8.36	-9.34	-5.03
Elasticity (Rec. Vary)	-11.71	-12.08	-6.41	-11.33	-11.86	-6.85

Table 10: Demand Estimation Results: Robustness to Time Horizon.
Notes: Estimates from maximum-likelihood estimation using data on 50,486 product pages spanning unique merchants between 2018-08-26 01:18:08 and 2020-03-25 23:39:41. The estimates are based on 64,809 sales for the 15m columns, 139,267 sales for the 1h columns, and 220,166 sales for the 24h columns. The reported coefficients measure the effect of price (normalized by the MSRP), the effect of an extra day of time until dispatch, the effect of a one percent increase in number of feedback, the effect of an offer being fulfilled by Amazon, and the mean utility of the outside option (relative to the inside option). λ is the nesting coefficient for sophisticated consumers and ρ is the fraction of consumers that are sophisticated. Elasticity (Rec. Fixed) refers to the average price elasticity of demand assuming that price does not influence consideration sets and Elasticity (Rec. Vary) refers to the said price elasticity under which it does. Elasticity (Rec. Var) is computed using the earlier estimates of the recommender system. All standard errors are clustered at the product level.

totals for sales on product pages – for instance, because they are considering entering on a particular product page. To fill this need, companies AMZScout and JungleScout offer "competitive intelligence" products that estimate the total sales for a product page using the product's sales rank within its top-level category. For each top-level category in our data, we obtained[38] sales estimates from both JungleScout and AMZScout for a log-spaced grid of sales ranks. We use this grid to linearly interpolate[39] (in log-log space) the total sales for each product in each month based on the average sales rank of the product in this month. We then average across months to estimate the total (monthly) sales for a product. All results in the main text utilize the AMZScout estimates obtained like this.

Alternatively, as a robustness check, we can use our sales data to estimate a relationship between log ranks and log sales. Denote by $A_{p,\tau}$ the number of sales for product p in month[40] τ. 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 τ (as in the main text). We will assume that log sales is affine in log sales ranks:[41]

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

This suggests estimating the following Poisson model of observed monthly sales against sales rank:[42]

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

However, we do not observe the fraction of sales $s_{p,j^o(p),\tau}$ that go to $j^o(p)$. Instead,

[38] We are grateful to Gutiérrez (2021) who shared the AMZScout data with us.

[39] For products with very high sales ranks, we have to extrapolate as both competitive intelligence companies round their daily sales estimates (so that we cannot distinguish e.g. 0.1 sales/day from 0.4 sales/day). We estimate the slope that governs the relationship between log sales ranks and log sales for the highest sales-rank interval unaffected by rounding. Then, we assume this slope continues to govern the relationship at higher ranks.

[40] We aggregate to the monthly level for this estimation; in the main text, τ refers to a day.

[41] 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.

[42] 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 replace $\ln\left(\frac{s_{p,j^o(p),\tau}}{1-s_{p,0,\tau}}\right)$ with a linear function of the (empirical) fraction of time that the offer is recommended, $\theta_3 \log(r_{p,j^o(p),\tau})$, a plausible proxy.

We further restrict attention to product-months 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 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 form an estimate of the number of sales on a product:

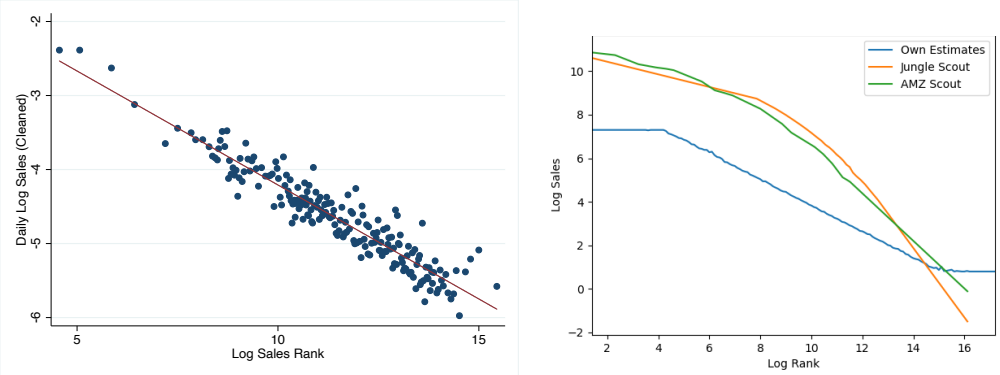
$$\hat{A}_p = \frac{1}{T_p} \sum_{\tau=1}^{T_p} \exp\left[\hat{\gamma}_0 + \hat{\gamma}_1 \ln\left(\text{rank}_{p,\tau}\right)\right].$$

Here, we substitute in the sales rank of a given product for each month, then set $r = 1$ to predict sales for a merchant on that product in this month. Thus, our prediction is made as if said merchant were the single merchant on a product; hence we recover an estimate of total sales. Finally, we average across months.

We illustrate our procedure in Figure 7. 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) for one. The linear model appears to fit the data well. Figure 7(b) compares our estimates to AMZScout and JungleScout. Both roughly agree with our estimates, though we slightly underestimate overall sales.

Finally, we need to translate monthly sales into market size for both methods. Doing so requires taking a stance on the number of consumers contemplating purchasing a product but ultimately deciding not to pull the trigger (i.e., choosing the outside option). We consider 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 (digitalcommerce360.com). Thus, we infer market size as $\hat{M}_p = \frac{1}{0.123}\hat{A}_p$.

As a robustness check, we run demand estimation with our own sales estimates and compare the results to our mainline estimates (which use AMZScout data)



(a) Relationship roughly linear. **(b)** Industry vs Own Estimates.

Figure 7: 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. This figure illustrates the procedure for the example category "Clothing, Shoes & Jewelry." As anticipated by prior literature, panel (a) illustrates that log sales are roughly linear in log sales ranks on the support of our data. Panel (b) compares our estimates to estimates from industry experts at AMZScout.com and JungleScout.com. We estimate slightly lower sales numbers at low sales ranks. Note that our estimates are constant in the tails because they are winsorized.

in Table 10. Other than a difference in the attractiveness of the outside option (necessary to reconcile the larger total sales estimates from AMZScout with our observed sales), we find no qualitative differences between the coefficient estimates under the two ways of specifying market size.

B.5 Maximum Likelihood Estimation of Consumer Choice

While exactly one alternative (or the outside option) is recommended for each market (p, τ), 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 in Section B.4 how we estimate the average market size M_p, i.e., the average number of customers that choose between inside and outside options on any given date τ and product page p. Given this number, we can form a pseudo-likelihood function[43] by assuming that the number of arrivals

[43]This function is only a pseudo-likelihood as our predicted number of arrivals M_p need not be an integer.

is distributed Poisson[44] with mean equal to M_p. Recall that $j^o(p)$ denotes the merchant for whom we observe sales on product page p. Furthermore, $a_{j^o(p),p,\tau}$ denotes the total number of sales for the observed merchant on product p and date τ in our data, and $s_{j^o(p),p,\tau}$ denotes her market share in our model. Thus, we can write the log-likelihood as:

$$\begin{aligned}\mathcal{L}(\alpha,\beta) &= \log\prod_p\prod_\tau \frac{exp(-s_{j^o(p),p,\tau}M_p)(s_{j^o(p),p,\tau}M_p)^{a_{j^o(p),p,\tau}}}{a_{j^o(p),p,\tau}!} \\ &\propto \sum_p\sum_\tau\Big[a_{j^o(p),p,\tau}\times\log(s_{j^o(p),p,\tau}) - s_{j^o(p),p,\tau}M_p\Big].\end{aligned}$$

C Details of Simulated Method of Moments

In this section, we detail the SMM procedure. We first describe how we solve the pricing and entry games; how we simulate a market; and how we aggregate markets to moments. Next, we describe concentrating out wholesale cost parameters, and smoothing out the SMM objective by importance sampling. Finally, we display measures of model fit.

C.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 other's types and treat price as a strategic variable. Thus, the first-order condition associated with seller j's choice of price is

$$\phi s_j(\mathbf{p},\mathbf{q}) + [\phi p_j - c_j]\frac{\partial s_j}{\partial p_j} = 0. \tag{1}$$

An equilibrium of the pricing game comprises prices $\mathbf{p}$ satisfying Equation (1) 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 *ζ-markup equation*. First, we decompose the Jacobian matrix of market shares:

[44] By contrast, if we assumed that exactly M_p consumers arrived each period, our model would not be able to rationalize markets with more sales than the average number of consumer arrivals.

$\frac{\partial \mathbf{s}}{\partial \mathbf{p}'} = \Lambda(\mathbf{p}) - \Gamma(\mathbf{p})$, 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. The ζ-markup equation is

$$\mathbf{p} = \phi^{-1}\mathbf{c} + \zeta(\mathbf{p}), \quad \zeta(\mathbf{p}) = \Lambda(\mathbf{p})^{-1}\Gamma(\mathbf{p})'(\phi\mathbf{p} - \mathbf{c}) - \Lambda(\mathbf{p})^{-1}\mathbf{s}(\mathbf{p}), \tag{2}$$

whenever $\Lambda(\mathbf{p})$ is nonsingular.[45]

Having described Equation (2), we chain iterations based on the ζ-markup and BLP-markup equations to find a fixed point. Iterating on the ζ-markup equation improves convergence relative to iterating on the BLP-markup equation alone.[46]

In our model, 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 adjusted recommendation system attractiveness, $q_j^r - \alpha^r c_j$, a lower starting price than the others. In (unreported) robustness checks, we find that varying the "privileged" entrant does not change our qualitative findings.

C.2 Solving the Entry Game

The information set[47] 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

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

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

$$\zeta(\mathbf{p}) = \Omega(\mathbf{p})(\phi\mathbf{p} - \mathbf{c}) + \Big(I - \Omega(\mathbf{p})\Big)\eta(\mathbf{p}); \quad \eta(\mathbf{p}) \triangleq -\left[\frac{\partial \mathbf{s}}{\partial \mathbf{p}'}\right]' \mathbf{s}(\mathbf{p}),$$

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

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

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

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.

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

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

C.3 Simulating a Market

We can now combine Subsections C.1 and C.2 to simulate market-level outcomes. We employ the algorithm presented in Figure 8b. 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 $\mathbf{q}^r_s \in \mathbb{R}^{|\mathcal{N}|}$. We employ the simulation draws and the algorithm of Subsection C.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 Subsection C.1.

C.4 Aggregating Markets to Moments

We now aggregate market-level outcomes to moments; our discussion follows Ackerberg (2009) adapted to our model. Fix a market m. Our model postulates a relationship f between (market-level) observables[48] $x_m = (A_m, R_m)$, unobservables u_m and outcomes y_m. In particular, at the true parameter vector θ_0 we have

$$y_m = f(x_m, u_m, \theta_0).$$

[48] Here, A refers to market size and R to the manufacturer's suggested retail price.

We begin by defining the outcomes included in y_m. Let the number of successful entrants J_m, the average price $\overline{p}_m$ and the standard deviation of log prices std_m on market m be given by

$$J_m := |\mathcal{J}_m| = \sum_{j \in \mathcal{N}_m} 1\{c_{jm} \leq c_m^*\}, \quad \overline{p}_m = \frac{1}{J_m} \sum_{j=1}^{J_m} p_{jm}, \text{ and } std_m = \sqrt{\frac{1}{J_m} \sum_{j=1}^{J_m} [\log(p_{jm}) - \overline{\log(p_m)}]^2}$$

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

$$y_{m,1} = \begin{pmatrix} J_m \\ J_m^2 \\ J_m \times A_m \\ J_m \times R_m \\ (y_{m,N}) \end{pmatrix}, \quad y_{m,N} = \left((1\{J_m = j\})_{j \neq \lfloor |\mathcal{N}|/2 \rfloor} \right) \qquad y_{m,2} = \begin{pmatrix} \overline{p}_m \\ \overline{p}_m \times R_m \\ std_m \end{pmatrix}.$$

If the data $\{x_m, y_m\}_{m=1}^M$ are generated by by our model at the true θ_0, then

$$\theta = \theta_0 \implies \mathbb{E}[y_m - \mathbb{E}[f(x_m, u_m, \theta)|x_m]|x_m] = 0. \tag{3}$$

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

$$F_m \sim \text{LogN}(\theta_0^F + x'_m \theta_x^F, \theta_\sigma^F) \text{ w.p. } \omega, \text{ and LogN}(\underline{\theta}_0^F, \underline{\theta}_\sigma^F) \text{ otherwise.}$$

Wholesale costs are drawn from

$$c_{jm} = (\theta_0^c + x'_t \theta_x^c)\epsilon_{jm}, \quad \epsilon_{jm} \sim \text{LogN}(0, \theta_\sigma^c).$$

Given some candidate $\theta = (\theta_0^F, \theta_x^F, \theta_\sigma^F, \omega, \underline{\theta}_0^F, \underline{\theta}_\sigma^F, \theta_0^c, \theta_x^c, \theta_\sigma^c)$, the conditional distribution of unobservables $p(u_m|x_m, \theta)$ is thus fully specified. In theory, we could use our model and the moment condition(s) in (3) to estimate θ. 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). Instead, we take simulation draws $(u_{m1}, \ldots, u_{mS}) \sim p(u_m|x_m, \theta)$ and approximate the expectation by averaging:

$$\widehat{\mathbb{E}f_m(\theta)} = \frac{1}{S}\sum_s f(x_m, u_{ms}, \theta).$$

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 $\hat{G}(\theta)$ close to zero:

$$\hat{\theta} = \arg\min_{\theta\in\Theta} Q(\theta) = \hat{G}(\theta)' W \hat{G}(\theta), \quad \hat{G}(\theta) = \frac{1}{M}\sum_m (y_i - \widehat{\mathbb{E}f_m(\theta)}),$$

for some weight matrix W.

C.5 Concentrating Out Wholesale Cost Parameters

We can partition $\theta = (\theta^F, \theta^c)$ where $\theta^F = (\theta_0^F, \theta_x^F, \theta_\sigma^F, \omega, \underline{\theta}_0^F, \underline{\theta}_\sigma^F)$ 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. We employ the algorithm presented in Figure 8a.

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 game.

for $p = 1, \ldots, P$ **do**
  for $s = 1, \ldots, S_F$ **do**
    simulate entry for (p,s)
$\theta^F = \theta_0^F$
while *dist* > *tol* **do**
  reweight sim data for θ^F
  dist = $\hat{G}(\theta^c, \theta^F)$
  update θ^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 8: 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 entry simulation.

C.6 Importance Sampling

Computationally, finding $\tilde{\theta}^F(\theta^c)$ requires a continuous objective. However, the entry game is discrete. 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 $\ln F_m = \theta_0^F + x_m' \theta_x^F + u_m^F$. As θ^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., $f(x_m, u_m, \theta) = \tilde{f}(F_m)$. We pursue the following simulation strategy: draw $F_{ms} \sim g(\cdot)$ where $g(\cdot)$ 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, we compute

$$\widetilde{\mathbb{E}f_m}(\theta) = \sum_s \tilde{f}(F_{ms}) \frac{p(F_{ms}|x_m, \theta)}{g(F_{ms}|x_m)}.$$

This simulator has two advantages. Firstly, the density for fixed costs we specified is smooth, ensuring that the simulated outcomes will smoothly depend on θ^F. Thus, 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 $\tilde{f}(F_{ms})$ once at the beginning of the procedure. As we vary θ^F, the outcomes employed do not vary but are 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 $g(\cdot|x_m)$ equal the density $p(\cdot|x_m, \theta^F_{prev})$ at the previous iteration's optimal value θ^F_{prev}. As in the original paper, this iteration converges extremely fast (typically in just three steps).

C.7 Results

We illustrate the fit of our SMM estimates in Figure 9. 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 9a illustrates the empirical distribution of mean price across products (in solid light orange) against the distribution implied by the model (outlined in blue). The model distribution matches its data analog extremely well. Similarly, Figure 9b 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. Many 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 captures across-product heterogeneity. For example, in 9c, 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, 9d illustrates that our model cannot capture much of the across-product variation in the number of entrants. This precludes us from making statements about, e.g., potentially heterogeneous effects of recommendation algorithm design.

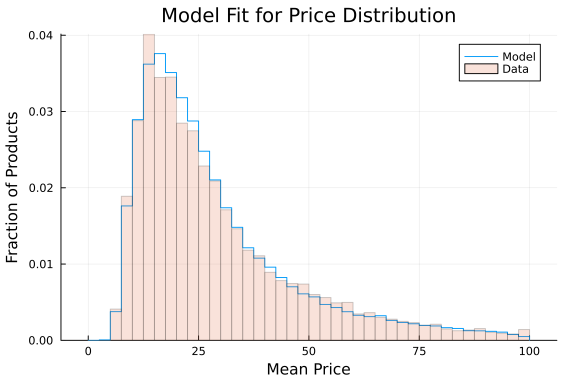
Why can our model not match the across-product variation in the number of entrants? On the one hand, we may just be missing good covariates that determine fixed costs. For instance, we have allowed fixed costs to vary by market size and suggested retail price. Perhaps, however, fixed cost varies across 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 the variation in (effective) fixed costs is just random? Concretely, it could be that discovering niches to enter is a fundamentally random process. As already discussed in the main text, turnover on Amazon is high. This suggests that there are a lot of unmodeled shocks to entry decisions. Under this line of argument, attempting to 'force' all dots in Figure 9d 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.

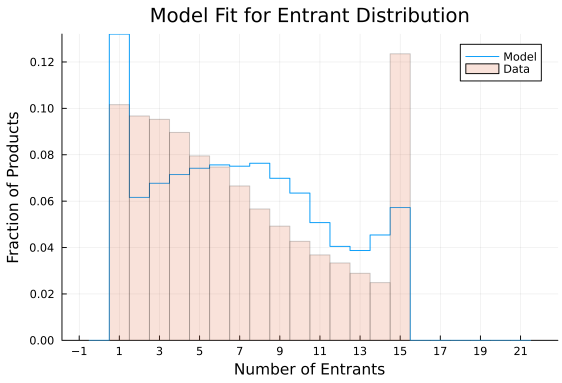
D Additional Counterfactuals

D.1 Platform Price Response to "Self-Preferencing"

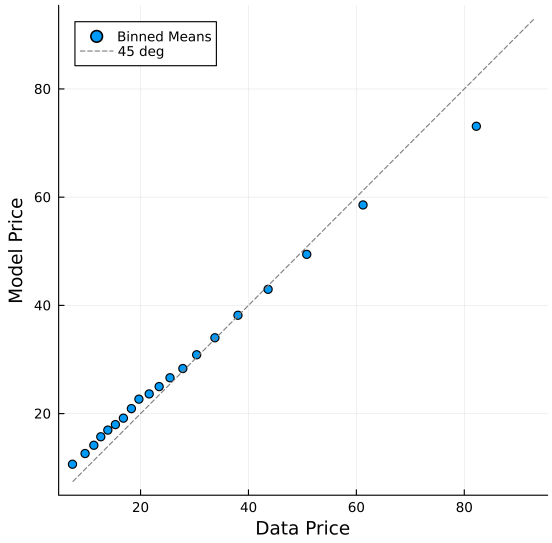
Our mainline results on the effect of "self-preferencing" do not account for the platform's pricing strategy. Indeed, we assume in the main text that the platform's prices and entry decisions remain fixed across all counterfactuals. Here, we relax this assumption. Rather than take a stance on the objective function that the platform is optimizing, we compute the price change required for our welfare results to flip. To do so, we progressively raise Amazon's prices in the factual world by 1.0% each time, continuing until we have reached the point at which the overall welfare effect of this platform advantage, net of the price increase, is (close to) zero. This point is passed once Amazon's prices are 2.0% higher in the observed scenario, relative to the counterfactual world. We exhibit the overall



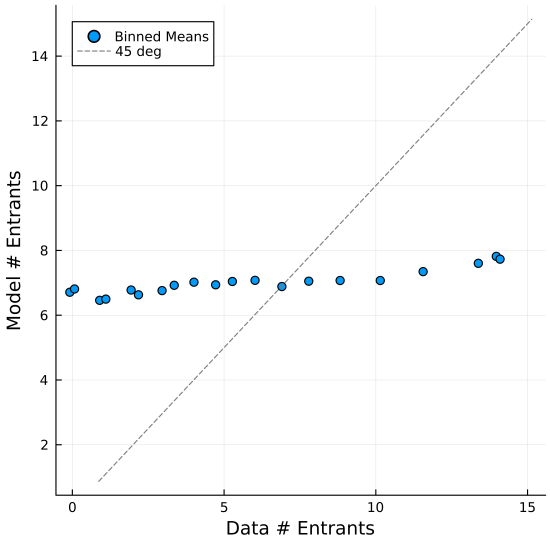
(a) Distribution of Mean Price.



(b) Distribution of # Entrants.



(c) Fit of Mean Price.



(d) Fit of # Entrants.

Figure 9: Fit of Distributions Implied By Marginal & Fixed Cost Parameters $\hat{\theta}$.
Notes: We exhibit the distributions implied by the values of the marginal and fixed cost parameters θ estimated as part of the SMM procedure and compares them to the corresponding empirical distributions. The fit for the price distribution is good: we not only match the distribution of prices, our model also predicts high prices for products for which we observe high prices in the data, and vice-versa. By contrast, the fit for the number of entrants is less good. While our model can rationalize the distribution of the number of entrants (though it struggles with the excess mass at the point where we censor the distribution), it cannot match the across-product variation in the number of entrants. This is because we do not observe any covariates highly predictive of across-product fixed costs heterogeneity. Hence, we restrict ourselves to examining aggregate effects.

	Short-Run	Medium-Run	Long-Run
Prices Adjust?	✗	✓	✓
Entry Decisions Adjust?	✗	✗	✓
Δ Consumer Surplus (CS)	-$501,002	-$447,047	-$444,739
Δ Consumer Surplus (Naive)	$51,530	$90,132	$92,986
Δ Consumer Surplus (Soph.)	-$552,532	-$537,179	-$537,726
Δ Producer Surplus (PS)	-$8,318	-$4,633	-$4,985
Δ Intermediation Fees	-$16,655	$25,852	$23,823
Δ Welfare (CS + PS)	-$509,320	-$451,680	-$449,725
Δ 3P Mean # Sales/Month	-0.03	0.02	0.02
Δ 3P Mean Price (% MSRP)		0.00%	-0.02%
Δ 3P Mean Min Price (% MSRP)		0.00%	0.00%
Δ 3P Mean Cost (% MSRP)			-0.02%
Δ 3P Mean # Entrants			-0.00

Table 11: Self-Preferencing is Welfare Negative if it Raises Platform Prices 2.0%.
Notes: This table displays differences in outcome variables, on our sample of 50,486 products, when the owner of the marketplace advantages its own offers in its recommender system and raises prices by 2.0%, relative to a counterfactual in which it does not. In the short run, sellers cannot change their prices or entry decisions. In the medium run, sellers may set prices optimally, but cannot change their entry decisions. Finally, the long-run counterfactual allows sellers to change both their prices and their entry decisions. Outcomes that cannot change (e.g., the mean number of entrants in the medium-run counterfactual) are omitted from the Table for clarity. 3P refers to 'Third-Party', i.e., it indicates that the outcome is computed using only non-Amazon offers.

effect of self-preferencing under this pricing difference in Table 11, where, as a reminder, the '3P Mean Price' line only reflects pricing-changes by third-party merchants (which barely adjust their price in response to Amazon's price increase partially because most products do not feature an Amazon offer.)

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 2.0% relative to the counterfactual world in which no self-preferencing occurs.

D.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

	Short-Run	Medium-Run	Long-Run
Prices Adjust?	✗	✓	✓
Entry Decisions Adjust?	✗	✗	✓
Δ Consumer Surplus (CS)	$3,130,577	-$4,441,222	-$3,107,138
Δ Consumer Surplus (Naive)	$3,130,577	$46,199	$517,806
Δ Consumer Surplus (Soph.)	$0	-$4,487,421	-$3,624,945
Δ Producer Surplus (PS)	$696,236	$917,636	$849,798
Δ Intermediation Fees	$1,431,849	$624,273	$724,115
Δ Welfare (CS + PS)	$3,826,813	-$3,523,586	-$2,257,340
Δ 3P Mean # Sales/Month	0.34	-0.49	-1.04
Δ 3P Mean Price (% MSRP)		2.04%	1.98%
Δ 3P Mean Min Price (% MSRP)		1.89%	0.89%
Δ 3P Mean Cost (% MSRP)			0.86%
Δ 3P Mean # Entrants			0.26

Table 12: There May Be Benefits to More Price Elastic Recommendations.
Notes: This table displays differences in outcome variables, on our sample of 50,486 products, when a recommender system is employed on our third-party marketplace, relative to a counterfactual in which the platform doubles its sensitivity to price. In the short run, sellers cannot change their prices or entry decisions. In the medium run, sellers may set prices optimally, but cannot change their entry decisions. Finally, the long-run counterfactual allows sellers to change both their prices and their entry decisions. Outcomes that cannot change (e.g., the mean number of entrants in the medium-run counterfactual) are omitted from the Table for clarity. 3P refers to 'Third-Party', i.e., it indicates that the outcome is computed using only non-Amazon offers.

the estimated price coefficient, and reporting in Table 12 the impact of moving from such a counterfactual algorithm to the factual algorithm actually employed.

Compared to a heavier emphasis on price, the algorithm employed in practice performs better in an A/B test but not in the long run. To begin with, the first column of Table 12 indicates that before taking pricing and entry decisions into account, the current algorithm benefits consumers, producers, and platform alike: its comparatively weaker (but still substantial) emphasis on price allows the platform to surface higher quality offers to consumers, which these are more likely to purchase. However, these benefits do not consider the effect of more elastic recommendations on pricing. Once we allow pricing to change, the current algorithm's weaker emphasis on price leads to 2.04% higher prices. Finally, there is slightly more entry under the current algorithm, though the difference of 0.26 entrants per market is small.