Two-Sided Markets Shaped by Platform-Guided Search

Kwok Hao Lee, Leon Musolff[1]

May 19, 2025

Abstract

This paper investigates concerns that vertically integrated platforms like Amazon steer demand towards their own offers via algorithmic prominence, potentially harming consumers. On Amazon, for each product, the Buybox prominence algorithm selects one seller to feature, influencing which offers consumers consider. Using novel Amazon sales and Buybox (prominence) data, we estimate a structural model capturing the effects of such algorithmic prominence on consumer choices, seller pricing, and entry. We find that the platform can indeed steer demand as 95% of consumers consider only the Buybox offer. The Buybox is highly price-elastic (-21), but skews towards Amazon's own offers, which are featured as frequently as observably similar offers priced 5% cheaper. Still, as consumers prefer these offers, this skew does not amount to self-preferencing in the sense of harming consumers: consumer surplus is roughly maximized at the estimated Amazon Buybox advantage, which balances higher prices against showing consumers their preferred offers.

Keywords: Two-Sided Markets, E-Commerce, Consumer Search, Self-Preferencing

[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. We thank the Editor, Francesco Trebbi, and three anonymous referees, for their invaluable feedback. Special thanks go to Jakub Kastl, Kate Ho, Nicholas Buchholz, Adam Kapor and Bo Honoré for invaluable guidance. We would also like to thank Bruno Baránek, Juan Camilo Castillo, Giovanni Compiani, Chris Conlon, Michael Dinerstein, Glenn Ellison, Chiara Farronato, Qiang Fu, Germán Gutiérrez, Andreas Haupt, Jean-François Houde, Ian Kaneshiro, Uri Miron, Devesh Raval, Imke Reimers, Tiffany Tsai, Raluca Ursu, Matthijs Wildenbeest, and Chuan Yu.

1 Introduction

Online platforms increasingly act not merely as passive intermediaries but actively shape consumer experiences. Through search, ranking, and prominence algorithms, platforms like Amazon direct consumer attention. These algorithms significantly impact market outcomes: by influencing which options consumers consider, platforms effectively steer demand, thereby shaping price competition, merchant entry, and ultimately consumer welfare. We call this ability to steer demand *intermediation power*.

In this paper, we study the implications of such intermediation power on Amazon Marketplace, where Amazon both operates the platform and competes as a seller. Vertically integrated platforms like Amazon may engage in self-preferencing – guiding consumers preferentially towards their own offers, potentially harming consumers (Stigler Committee on Digital Platforms 2019). This concern has drawn scrutiny from antitrust agencies (Federal Trade Commission and State Attorneys General 2023) and regulators (European Commission 2022; U.S. Congress 2021; U.S. Senate Judiciary Committee 2020). Besides potential static harm from not showing consumers their most preferred offers, proposed theories of harm from self-preferencing include a possible stifling of merchant entry and associated reduction in options available to consumers (Klingler et al. 2020).

Our primary research question is whether Amazon's Buybox prominence algorithm is self-preferencing in the sense of harming consumers. To this end, we build and estimate 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 prominence assignments (which steer consumers), pricing, and entry. We combine this model with unique proprietary data linking sales and prominence assignments on Amazon, a setting of prime regulatory concern. Our counterfactuals speak to various theories of harm from self-preferencing and the platform's tradeoff between attracting merchants and fostering competition.

Platforms commonly set defaults for consumers through their search (Yu 2024), ranking (Ursu 2018), and prominence algorithms (Johnson, Rhodes, and Wildenbeest 2023). Economists have increasingly become aware that these defaults can influence consumer choice. We thus model consumers as having one of two kinds of consideration sets (Ben-Akiva and Boccara 1995; Manski 1977): either they consider all available options, or they only consider the prominent offer and the outside option. Through its influence on consideration sets, the platform's prominence algorithm affects seller pricing (as in, e.g., Dinerstein et al. 2018), as well as *how many* and *which* firms enter.

We estimate this model using high-frequency Amazon data via maximum likelihood (prominence algorithm and demand) and GMM (pricing and entry). Variation in offer characteristics (and corresponding prominence assignments) identifies how Amazon decides which offers to make prominent. Similarly, variation in choice sets and observed sales identify consumer demand parameters. The fraction of consumers who consider all offers (as opposed to only the prominent offer) is identified by the difference in sales between offers eligible for prominence and those not eligible. The estimated demand structure and variation in merchant prices identify marginal costs. Finally, the model of variable profits implied by the estimated demand system and variation in market size identify fixed costs.

We find that Amazon can steer demand and indeed has a large influence over it as 95% of consumers consider only the Buybox offer. The Buybox is highly price-elastic (-21), but skews towards Amazon's own offers, which are featured as frequently as observably similar offers priced 5% cheaper.

After recovering the prominence algorithm, demand, and cost parameters, we simulate market outcomes under alternative prominence algorithms, proceeding in two stages: first with fixed prices and entry (short run), then allowing price and entry adjustments (long run). We consider counterfactuals that examine (i) the value of search guidance, (ii) how the price sensitivity of prominence affects consumer surplus, and (iii) self-preferencing, i.e., whether the frequency with which Amazon promotes its own offers is consumer surplus maximizing.

First, we evaluate the impact of search guidance by comparing the status quo to a benchmark in which unsophisticated consumers only consider the cheapest offer. While stylized, this counterfactual illustrates how search guidance shapes market outcomes. We find that the algorithm increases consumer surplus in the short run, mostly because of better matching of consumers to high-quality offers. Relative to our benchmark, the algorithm also increases producer surplus in the short run, which leads to additional entry and lower prices in the long run. Although these effects are modest, together they reinforce the increase in consumer surplus.

Second, we further investigate the prominence algorithm's price sensitivity, addressing concerns that the platform, which makes revenue from *ad valorem* intermediation fees, may be insufficiently incentivized to foster competition. Contrary to such concerns, we find that increasing the algorithm's price sensitivity would lower consumer surplus. Indeed, our evaluation suggests that the platform's price sensitivity is approximately optimal from a consumer surplus perspective.

Third, we investigate the question at the core of our paper: is Amazon self-preferencing in the sense of harming consumers? To answer this question, we explore how the frequency at which Amazon currently promotes its own offers compares to that which would maximize consumer surplus. While more frequent promotion of Amazon offers leads it to raise prices, it also allows more consumers to consider Amazon's offers, which they strongly prefer. We find that at the estimated promotion frequency, these two effects approximately balance, so that Amazon is not self-preferencing in the sense of harming consumers: increasing the Amazon promotion frequency would lead to consumer surplus losses from increased prices, but lowering it would lead to losses from worse matching of consumers to their preferred offers. The current promotion frequency is thus approximately consumer surplus maximizing.

2 Literature Review

This paper speaks to three strands of literature in industrial organization and the economics of digitization.

First, we contribute to a large literature on price competition and search frictions in online markets. Despite theoretically small search costs, online markets still exhibit 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; Schneider 2016; Stigler 1961) and consider only a subset of available alternatives (Goeree 2008). Given 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). Efficient search technologies can lead to high offer-level elasticities, prompting retailers to obfuscate product characteristics to hinder comparison and raise profits (Ellison and Ellison 2009). Finally, personalized prominence as in recommender systems can also affect pricing incentives (Calvano et al. 2025). While addressing similar platform-design issues, this literature rarely models entry or a firm competing on its own platform.

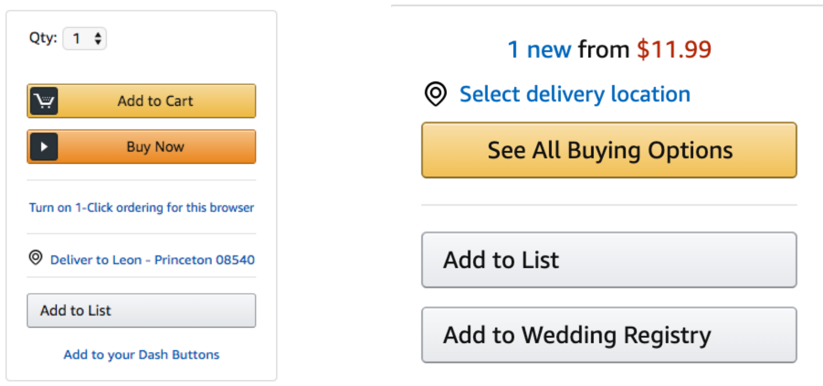
Second, we contribute to the literature on the downstream conduct of vertically integrated firms. In traditional markets, concerns include foreclosure (Crawford et al. 2018), raising prices of rivals' products (Luco and Marshall 2020) or altering products to disadvantage competitors (McManus et al. 2020). In digital markets, the focus shifts to steering, with self-preferencing defined as "platform [prominence assignments that] deliver outcomes below [a] frontier that maximizes a weighted sum of seller and consumer surplus" (Reimers and Waldfogel 2023). Such steering may negatively affect consumers through higher commissions and prices (e.g., Hagiu, Teh, and Wright 2022; Teh and Wright 2022).

Third, spurred partly by Khan (2016), our empirical setting – the Amazon marketplace – has emerged as a central focus for studies of price competition and search on platform marketplaces (Ciotti and Madio 2023; Crawford et al. 2022; Gutiérrez 2021; Lam 2021), including research on steering (Chen and Tsai 2023; Farronato, Fradkin, and MacKay 2025; Hartzell and Haupt 2025; Raval 2022; Waldfogel 2024). Etro (2022) provides a survey. Our paper establishes that algorithmic steering exists and is of first-order relevance to consumer welfare.

We highlight two particularly relevant studies. Farronato, Fradkin, and MacKay (2025) experimentally alter Amazon brand visibility in search results, finding consumer surplus would fall if these brands were removed. We, however, investigate a distinct steering dimension (Buybox prominence), which has garnered separate regulatory scrutiny (U.S. Senate Judiciary Committee 2020). Raval (2022) also investigates the Buybox, finding steering towards Amazon. Our work, using sales data, analyzes not just the presence but also the welfare effects of this steering.

3 Who does the Amazon Buybox feature?

Our empirical setting is Amazon Marketplace, a platform permitting third parties to list offers alongside Amazon's. In 2020, about 1.9 million merchants utilized this opportunity (aboutamazon.com). Their listings were responsible for 60% of sales on this marketplace, yielding estimated merchant profits of $25 billion (aboutamazon.com). However, with 28% of purchases completed within three minutes, consumers use little time to explore their options (aboutamazon.com).



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

Figure 1: The Buybox.

Notes: The "Amazon Buybox". 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 featured.

This speed suggests that search, ranking, and prominence algorithms play a role in shaping consumers' choices.

E-commerce platforms typically employ multiple such algorithms. However, investigating algorithm behavior is challenging when products differ on unobserved dimensions. We exploit a unique feature of our setting to sidestep this issue: because Amazon requires sellers to list their offers on the correct product page, a specific product often has multiple offers.

When multiple offers are present, the platform automatically designates *at most one* offer as prominent, i.e., featured in the "Buybox" (Figure 1). This offer does not vary across consumers.[2] Often, Amazon "rotates" prominence between sellers of sufficiently high quality to "avoid stock-outs" and encourage sellers to "delight their customers" and "get good feedback" (feedvisor.com). Prominence is critical for sellers, as the "vast majority of sales" go through the "Buybox" (europa.eu), with estimates from 80% (U.S. Senate Judiciary Committee 2020) to 98% (Federal Trade Commission et al. 2023, p. 29).

Econometrically, 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 decided on this product, consumers must choose between various offers. For

[2] In rare situations, two offers are prominent: one for Prime customers, and one for non-Prime customers. We drop these rare cases.

instance, the merchant "Zappos" offers these shoes for US$57.68 and will deliver them via Fulfillment by Amazon (FBA), a "fulfillment service that allows businesses to use Amazon to store, pick, pack, and ship customer orders" (amazon.com). Alternatively, "BHFO" charges only US$54.99 but does not offer FBA. Below, we examine the influence of prominence on how consumers choose between such offers. Crucially, in this choice, all options share all product characteristics – obviating the need for complex demand models to explain substitution.

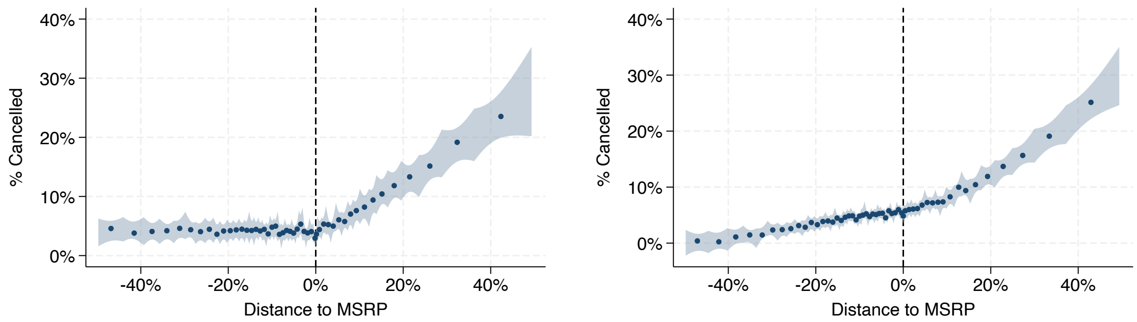
Amazon Marketplace is a key setting to evaluate antitrust concerns. Economists worry that "defaults can direct a consumer to the choice that is most profitable to the platform" (Stigler Committee on Digital Platforms 2019). Indeed, sellers complain that competing with Amazon is challenging, as the platform frequently assigns the Buybox to itself. A Senate investigation concluded that it "can give itself favorable treatment [and] has done so through its control over the Buy Box" (U.S. Senate Judiciary Committee 2020). The investigation sparked the introduction of a draft bill in Congress prohibiting self-preferencing (U.S. Congress 2021).

Recent allegations also claim Amazon modified its search algorithms to favor its own products (wsj.com), though evidence is mixed (Farronato, Fradkin, and MacKay 2023, 2025; Waldfogel 2024). These claims suggest seller concerns about the Buybox may be justified. However, Amazon's general counsel counters 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).

3.1 Data

We procured extensive, high-frequency data that includes prices, 19,409,013 prominence assignments, and 803,849 sales on 47,403 products sold by 63,620 merchants from 2018-08-26 until 2020-03-23. 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. For each monitored product, 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 assigns prominence to *just after* any such



(a) Unresidualized **(b)** Residualized

Figure 2: Evidence of Canceled Prominence (Binned Scatter Plots).
Notes: This binscatter illustrates the relationship between canceled prominence and the price of the cheapest offer on a product relative to MSRP; dots are binned averages and the grey shadow provides a 95% confidence band. The underlying observations are product-date pairs for which the lowest-priced offer is priced between 50% below and 50% above MSRP. The left panel shows the unresidualized relationship; the right panel depicts the relationship after residualizing on product fixed effects. In both panels, as the cheapest offer's markup over MSRP increases, canceled prominence becomes more likely.

change. However, prominence assignments may change frequently, and we receive no notification when the only change is which offer is assigned prominence.

We further observe, via a different API, all sales for *a single merchant* in each market — the one using the repricing company's services. These data come with their own challenges (e.g., having to proxy for the market size), which we address with our model and estimation strategy in Sections 4 and 5.[3]

3.2 Canceled Prominence

We first investigate when the platform cancels prominence, leaving the Buybox empty as in Figure 1b. This helps us understand the platform's incentives in shaping the prominence algorithm.

While an empty Buybox may lower sales, the threat of canceled prominence can discipline prices. Indeed, before 2019, Amazon's US marketplace required its third-party sellers to offer lower prices than on other platforms (theverge.com). These "most-favored nations clauses" (MFN) were deemed anticompetitive by European antitrust agencies, leading Amazon to drop them in Europe by 2013 (German Competition Authority 2013, 2015; UK Competition Authority 2013). Following similar pressure, Amazon also dropped the US MFN clauses by 2019.

[3]Yet, to our knowledge, no past papers employ sales data across many merchants to estimate demand on Amazon Marketplace. Instead, employing sales ranks as a proxy has become common.

	Fraction of Prominence Assignments		
	Overall	FBA	Amazon
Amazon Offer Exists	19.25%	19.83%	100.00%
FBA Offer Exists	97.07%	100.00%	100.00%
Prominent Offer Is ...			
Lowest Priced	40.60%	40.24%	45.27%
Second Lowest Priced	26.42%	26.73%	30.05%
Highest Feedback Count	29.32%	29.52%	66.67%
Highest Feedback Rating	19.17%	19.15%	8.90%
Fastest Shipping	25.71%	25.74%	31.22%
Lowest Priced FBA	50.31%	51.83%	64.96%
FBA	85.98%	88.57%	97.18%
Amazon	12.83%	13.22%	66.67%
Lowest Fastest Shipping	51.92%	52.36%	65.46%

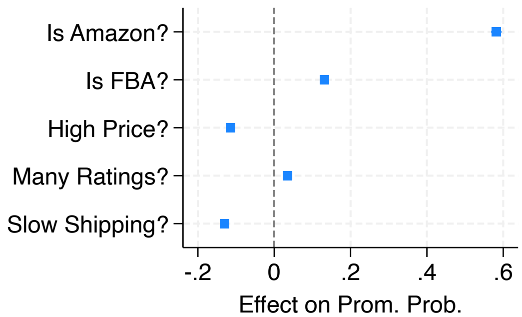
Table 1: Determinants of Prominence.

Notes: Each entry in this table gives the fraction of prominence assignments 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.

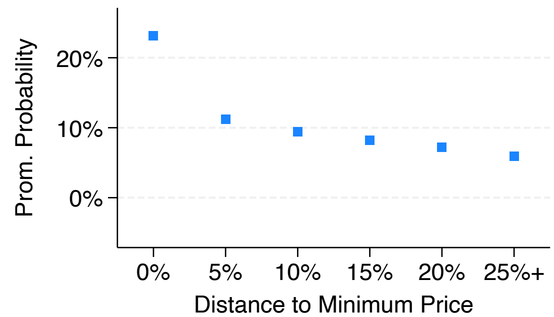
Nonetheless, canceled prominence may achieve similar effects to explicit MFN. To demonstrate, we combine data on the price of the lowest-priced offer on product j at date τ with data on Buybox ownership and information on the manufacturer's suggested retail price (MSRP), which we use as a proxy for the price at which a product could be obtained outside of Amazon. Figure 2 displays a binned scatter plot of Buybox cancellation versus the percentage distance to MSRP. As expected, the cancellation fraction increases with the cheapest offer's price. However, the increase is sharper if the offer price exceeds MSRP.

3.3 Prominence Descriptives

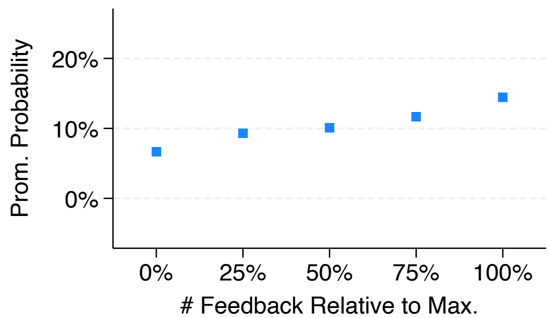
How does the platform choose which offer to feature? Table 1 investigates determinants of prominence. For instance, the prominent offer is the cheapest in 41% of cases. However, Fulfilled by Amazon (FBA) offers are given prominence in 86% of cases, possibly because they dispatch much faster. Most interestingly, on products where Amazon has an offer, its offer is featured 67% of the time.



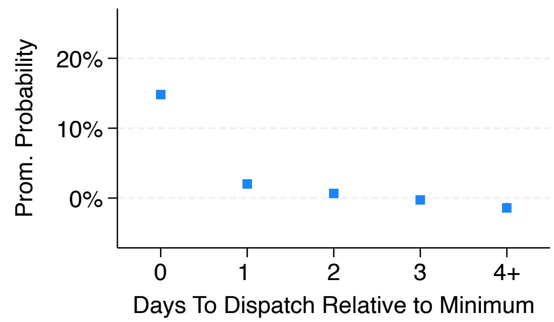
(a) Binarized Characteristics.



(b) Price.



(c) Feedback Count.



(d) Shipping Time.

Figure 3: More Determinants of Prominence.

Notes: Panel (a) shows the result of regressing a prominence dummy on binary characteristics (continuous characteristics were binarized by cutting at the median) without any further controls. Panels (b)-(d) show results from regressing the same prominence dummy on various offer characteristics – normalized by the product's best offer's value – while controlling for offer fixed effects. Standard errors are clustered at the product level (but confidence intervals are too small to be visible).

Figure 3 provides more detail on how offer characteristics matter. The top-left panel shows coefficients from bivariate regressions of a prominence dummy on (binarized) offer characteristics; continuous characteristics are split at the median. This figure shows that being sold by Amazon most strongly predicts prominence; however, as the underlying regressions are bivariate, this could reflect Amazon's offers being more competitively priced or of higher quality. FBA offers are also more likely featured, while high price, low feedback count, and slow dispatch times all correlate with lower prominence likelihood.

The remaining panels further investigate these three continuous characteristics. Acknowledging the prominence decision's discrete-choice nature, they normalize each characteristic relative to the product's best offer (e.g., expressing price as a percentage distance relative to the cheapest offer for the same product). These

	Norm. Sales	Norm. Sales/Market Size			
	(1)	(2)	(3)	(4)	(5)
Amazon Offer?	-38.72***	-26.81***	-28.10***	-23.61***	-11.50***
	(9.978)	(1.567)	(1.687)	(1.590)	(1.931)
# Comp. Merchants			-0.761***	-1.591***	-1.190***
			(0.176)	(0.176)	(0.155)
Divide by Mkt. Size	No	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes
Offer Fixed Effects	No	No	No	No	Yes
Observations	19,409,013	19,409,013	19,409,013	19,409,013	19,408,472

Table 2: Effect of Amazon's Presence on Third-Party Sales.

Notes: We regress normalized sales for the observed merchant on a dummy for whether there is an Amazon offer on the same market (i.e., product page and time.) Sales are normalized by dividing by their mean and multiplying by 100; regression coefficients can be interpreted as percentage point changes in sales. The first column uses raw sales data (showing that the effect is not due to our market size proxy); other columns use sales divided by market size (increasing power.) The # *Comp. Merchants* row counts the number of merchants pricing within 5% of the cheapest offer (for comparison, on average Amazon prices 7% above the cheapest offer.) *Controls* refer to price divided by MSRP, whether the offer is fulfilled by Amazon, shipping time, and the logarithm of the number of reviews. The effect of an Amazon offer on sales is negative, large, and statistically significant, even when controlling for the number of competitive merchants.

panels report coefficients from regressing a prominence dummy on evenly spaced bins of the normalized characteristic, controlling for offer fixed effects. Price is an important determinant, but even offers 20% above the minimum are still featured 7% of the time. Similarly, feedback count and especially dispatch time matter but are not solely decisive. The dispatch time effect centers almost entirely on whether shipping is instant. Crucially, these results do not control for other characteristics when looking at the impact of any given characteristic – in other words, the coefficients should not be understood as causal effects. Below, we estimate a discrete-choice model to disentangle the various factors influencing prominence, and provide assumptions under which the results of that model can be interpreted as causal.

Having shown that Amazon frequently features its own offers, we now turn to demand: is it true that Amazon offers are preferred by consumers? To investigate this question without access to sales data for Amazon offers – Amazon is not a client of the repricing company that is our data source – we turn to the idea that a strong preference for Amazon offers would be reflected in reduced sales for the observed merchants whenever Amazon is present on a market. We indeed find such a pattern in Table 2, which provides model-free evidence that Amazon's presence casts a large shadow on the observed merchant's sales by regressing the observed merchant's normalized sales on an Amazon presence dummy (the normalization allows interpreting the coefficient as a percentage sales change). We caution that this shadow could be caused either by a consumer preference for Amazon offers or by consumers only considering prominent offers (which we have just seen may favor Amazon offers). To disentangle these two possibilities, we next build a model of how consumers choose in the presence of steering by a prominence algorithm.

4 Model

To disentangle consumer preferences from the influence of prominence on choices, and to assess the welfare impact of alternative prominence algorithm designs, we build a model of consumer choice and within-platform price competition and entry. This model lets us examine whether designers of online marketplaces derive intermediation power from choosing their prominence algorithms. Our approach focuses on the platform's ability to guide search by influencing customers' consideration sets (Ben-Akiva and Boccara 1995; Manski 1977).

This section proceeds as follows. First, we specify demand: consumers have nested logit preferences with endogenous consideration sets. Second, we consider how the platform's prominence algorithm, also a nested logit, maps offer characteristics to a prominence assignment. Finally, we specify supply. On each market, potential sellers observe their own marginal costs and market-level fixed costs, make entry decisions, and then compete in a standard Bertrand-Nash pricing game. Crucially, sellers anticipate that demand depends on prominence, which hence shapes their incentives to enter and compete.

4.1 Consumer choice

Demand follows a standard nested-logit framework except that consumers form endogenous consideration sets (Dinerstein et al. 2018; Goeree 2008).

There are two types of consumers. A fraction ρ are "sophisticated." Sophisticated consumers ignore the prominence assignment and evaluate all available options $\mathcal{J} \cup \{0\}$. The remaining $1-\rho$ consumers are "unsophisticated." Unsophisticated consumers consider only the prominent offer $j^r \in \mathcal{J}$ and the outside option. Thus, the prominence algorithm chooses which offer the unsophisticated consumers evaluate.[4] If the algorithm chooses not to make any offer prominent, unsophisticated consumers take their outside option.

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}:

$$\delta_{jt} = \mathbf{x}'_{jt}\beta - \alpha_t p_{jt} + \xi_{jt}. \tag{1}$$

When discussing identification below, we impose further restrictions on ξ_{jt}: our main estimates utilize $\xi_{jt} \equiv 0$, but we also explore robustness to a more flexible specification. We also normalize the mean utility of the outside option to zero: $\delta_{0t} = 0$.

By contrast to the usual setup where markets are repeated observations of the same market over time (Berry, Levinsohn, and Pakes 1995) or different geographic markets with nearly identical alternatives (Nevo 2001), in our setting markets are distinct product pages. Hence, the alternatives offered differ markedly across markets, and so could price sensitivity. We assume that consumers evaluate price relative to MSRP. If R_t is the MSRP in market t, we set $\alpha_t = \alpha / R_t$, reflecting the idea that a \$0.50 price difference matters for a \$10 pen but not a \$1,000 laptop.

The flipside of alternatives differing markedly *across* markets is that *within* markets, the inside options available to consumers are very similar: 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 buying other products on the platform, buying the same product on some other platform or offline, and 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 alternative j in market t is then

$$v_{ijt} = \delta_{jt} + \zeta_{ig(j)t} + (1 - \lambda)\epsilon_{ijt}, \tag{2}$$

[4] Note that, 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-prominent offers.

where ϵ_{ijt} is distributed i.i.d. Type-1 Extreme Value, and $\zeta_{ig(j)t}$ is common to all options in the same nest. 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, yielding a nested-logit model (Cardell 1997; McFadden 1978).

Consumers choose the utility-maximizing option in their consideration sets. Therefore, given a prominent offer j^r, the probability that a consumer chooses product $j \neq 0$ in market t depends on a prominence dummy $y^r_{jt} \equiv \mathbf{1}\{j = j^r_t\}$:

$$\begin{aligned} s_{jt}(y^r_{jt};\theta^d) \quad &= \quad \rho \times \frac{\left[\sum_{k\in\mathcal{J}_t}\exp(\delta_{kt}/\lambda)\right]^\lambda}{1+\left[\sum_{k\in\mathcal{J}_t}\exp(\delta_{kt}/\lambda)\right]^\lambda} \times \frac{\exp(\delta_{jt}/\lambda)}{\sum_{k\in\mathcal{J}_t}\exp(\delta_{kt}/\lambda)} \\ &\quad\; +(1-\rho)\times y^r_{jt} \times \frac{\exp(\delta_{jt})}{1+\exp(\delta_{jt})}, \end{aligned} \tag{3}$$

where $\theta^d = (\beta, \alpha, \lambda)$ collects the parameters of the demand model.

We assume that the number of customers making a choice in market t (i.e., the number of arrivals) has a Poisson distribution with mean M_t.[5] By the thinning property of the Poisson distribution (Blitzstein and Hwang 2014, Theorem 13.12.13), the number of customers choosing j on market t is then also Poisson with mean $M_t \times s_{jt}(y^r_{jt};\theta^d)$.

4.2 Prominence Algorithm

A *prominence algorithm* maps offer characteristics to a prominence assignment $j^r \in \mathcal{J}$ or – if prominence is cancelled as in Section 3.2 – the null assignment $j^r = \{0\}$. While, in general, prominence may vary by consumer, this is not the case in our empirical application. Hence, we model the prominence algorithm as solving exactly one discrete choice problem for each market.

The inner workings of the prominence algorithm are not known. But whether it is a machine learning algorithm that maximizes an objective such as total sales or a simpler rule-based routine, what matters to us is how the prominence algorithm shapes demand, pricing, and entry. To this end, the important question is how often sellers can expect their offers to be promoted, and how this depends on their own and their rivals' characteristics, including price and whether they are

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

competing with Amazon. As with consumer choice, we capture this dependence parsimoniously with a nested-logit model. To the extent that the algorithm trades off various goals (e.g. sales profits vs seller surplus), counterfactuals should be understood as altering the weight given to these goals (e.g. lowering the weight on sales profits could correspond to lowering the Amazon prominence advantage).

The mean utility δ^r_{jt} of each inside alternative includes observable characteristics $\mathbf{x}_{jt}$, price p_{jt} and the econometrician-unobservable quality ξ^r_{jt}:

$$\delta^r_{jt} = \mathbf{x}'_{jt}\beta^r_t - \alpha^r \frac{p_{jt}}{R_t} + \xi^r_{jt}. \tag{4}$$

As with demand, when discussing identification below, we impose further restrictions on ξ^r_{jt}: our main estimates utilize $\xi^r_{jt} \equiv 0$, but we also explore robustness to a more flexible specification. We also normalize the mean utility of the outside option to zero: $\delta_{0t} = 0$.

Whether due to deliberate randomization or mistakes in evaluation, we assume prominence decisions are based on a shocked version v^r_{jt} of this mean utility. In particular, shocks follow a nested logit structure like with consumer choice.[6] Like before, all inside alternatives share a common nest, and the outside option is in a separate nest; unlike before, this nesting here reflects the fact that cancelling prominence is likely to lower intermediation fee revenue, and hence is qualitatively different from just assigning prominence to a different offer. As the shock structure is perfectly analogous to consumer choice, we will not repeat the details here. Instead, writing the nesting coefficient of the Buybox as λ^r, we skip ahead to the implied probability of prominence:

$$r_{jt}(\theta^r) = \frac{\left[\sum_{k\in\mathcal{J}_t} \exp(\delta^r_{kt}/\lambda^r)\right]^{\lambda^r}}{1 + \left[\sum_{k\in\mathcal{J}_t} \exp(\delta^r_{kt}/\lambda^r)\right]^{\lambda^r}} \times \frac{\exp(\delta^r_{jt}/\lambda^r)}{\sum_{k\in\mathcal{J}_t} \exp(\delta^r_{kt}/\lambda^r)}, \tag{5}$$

where $\theta^r = (\beta^r, \alpha^r, \lambda^r)$ collects the parameters of the prominence model.

[6]One might assume Amazon's prominence algorithm is *deterministic*, implying the variance of the logit shock ϵ is (near) zero. However, industry sources and our web scraping indicate stochasticity via "Buybox rotations": the platform shares arrivals among similarly priced merchants. Therefore, ϵ contains said rotations and other idiosyncratic factors affecting prominence.

4.3 Firm Choices

Finally, we specify a model of entry and price competition. Each market $t \in \mathcal{T}$ is associated with a fixed cost F_t and a set of potential entrants $j \in \mathcal{N}_t$.

Fixing a market t, each putative seller j draws a type $\omega_j \sim G(\cdot)$[7] and chooses a pair (χ_j, p_j) consisting of an entry and a pricing strategy. The entry rule $\chi_j : \mathcal{I}_j \to \{0, 1\}$ maps a player's pre-entry information set $\mathcal{I}_j$ into her entry decision, whereas the pricing rule $p_j : \mathcal{I}' \to \mathbb{R}^+$ maps the post-entry information set $\mathcal{I}'$ into the price the seller will charge.

As merchants receive only a share ϕ of the revenue they generate due to intermediation fees, and assuming constant marginal costs, seller j's payoff is

$$\pi_j(\boldsymbol{\omega}, \boldsymbol{\chi}, \mathbf{p}) = \sum_{t \in \{\tilde{t} | j \in \mathcal{N}_{\tilde{t}}\}} \chi_{jt} \times \left[(\phi p_{jt} - c_{jt})\, s_{jt}(r_{jt}(\boldsymbol{\omega}, \mathbf{p}), \boldsymbol{\omega}, \mathbf{p}) M_t - F_t \right], \tag{6}$$

where s_{jt} maps the vectors of types and prices into demand and M_t denotes expected monthly arrivals in market t. Our solution concept is Bayesian Nash Equilibrium.

We proceed by stating several assumptions for tractability:

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

Assumption 2 (Full Information Pricing)**.** *After entry, each seller knows the identities, cost draws, and quality draws of its opponents when playing the pricing game:* $\mathcal{I}'_j \equiv \mathcal{I}' = (\boldsymbol{\omega}, F)$.

Assumption 3 (Platform Offer Maximizes Short-Run Profits)**.** *On each market, the platform's decision whether to list its first-party offer is taken as given. The price of the platform offer is set market-by-market to maximize the profits accruing to the said offer.*

Assumption 4 (Blind Entry)**.** *Before entry, each seller observes only its own marginal cost draw and the common fixed cost of entry:* $\mathcal{I}_j = (c_j, F)$.

Assumption 5 (Symmetric Entry)**.** $\chi^*_j(c, F) = \chi^*_k(c, F)$ *for all* $j, k \in \mathcal{N}$ *and* $c, F \in \mathbb{R}^+$.

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

The separable-markets assumption allows us to solve for equilibrium market-by-market, easing the computational burden. In our empirical application, firms are small and rarely specialize in related products, so cross-market cannibalization is of secondary importance.

Full-information pricing is standard: successful entrants play a Bertrand-Nash game (Anderson, De Palma, and Thisse 1992) and internalize the effect of prominence when setting prices,

$$\frac{\partial s_j}{\partial p_j} = \rho \frac{\partial s_j^{sophisticated}}{\partial p_j} + (1-\rho)\left[\frac{\partial r_j}{\partial p_j} s_j^{unsophisticated} + r_j \frac{\partial s_j^{unsophisticated}}{\partial p_j}\right]. \tag{7}$$

By influencing $\frac{\partial r_j}{\partial p_j}$, the prominence algorithm therefore partially determines the price elasticity sellers face — exactly where the platform's power to guide consumer search manifests in our model.

As for the platform offer maximizing short-run profits, we make this assumption for tractability and because it is an important benchmark for our analysis. Appendix G.2 explores how results change if Amazon's prices are fixed in the counterfactuals.

Given a selection rule for the pricing equilibrium,[8] Assumptions 4 and 5 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 potential entrant independently decides whether to enter after observing information set $\mathcal{I}_j = \{c_j, F\}$. This assumption consists of two parts. Firstly, in the spirit of Roberts and Sweeting (2013) and to avoid the classic equilibrium multiplicity problems, we deprive sellers of knowledge about offers *other* than their own.[9] Secondly, before entry, each firm's only information about its type is its cost. Because (we estimate that) price is the most important determinant of prominence, this assumption will

[8]In our simulations, pricing equilibria always exist but are not necessarily unique as firms differentiate by targeting sophisticated or unsophisticated consumers. When multiple equilibria exist, we select one by starting our price fixed-point search with a lower price for the seller possessing the highest prominence algorithm attractiveness. Further details are in Appendix C.1.

[9]We make this assumption to avoid equilibrium multiplicity, as is common in games of strategic substitutes (e.g., Tamer 2003). We believe this assumption is innocuous because most sellers in our sample are small and anonymous; they are unaware of their competitors before entry. Even if 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.

only slightly affect our results.

Each firm enters 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^*. Conditional on a selection rule for pricing equilibria,

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 marginal cost cost weakly below c^*.

5 Identification and Estimation

We now discuss identification and estimation of $\theta = (\theta^r, \theta^d, \theta^s)$ where $\theta^r = (\beta^r, \alpha^r, \lambda^r)$ govern the prominence algorithm; $\theta^d = (\beta, \alpha, \lambda, \rho)$ govern demand; and $\theta^s = (\theta^F, \theta^c)$ govern the fixed F and marginal cost c components of supply. We estimate θ^r and θ^d via maximum likelihood, and θ^s via GMM.

We estimate our model using data covering 19,409,013 prominence assignments and 803,849 sales across 47,403 products and 63,620 merchants from 2018-08-26 to 2020-03-23. Recall that we observe individual-level sales data and discrete choice data for prominence assignments for each market t, which we define to be a "product (web)page"-notification time pair (w, τ) for the purpose of prominence and demand estimation.[10] 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.

[10]When considering pricing and entry, we aggregate to a "product page"-month.

5.1 The Prominence Algorithm

5.1.1 Main Model: No Unobserved Quality

We make the identifying assumption of no unobserved prominence quality, i.e.,

Assumption 6. *There is no unobserved prominence quality:* $\xi^r_{jt} \equiv 0$.

Intuitively, this assumption reflects the fact that alternatives in our model are offers, and markets are product-webpages. Hence, product characteristics (like color) cancel when comparing across alternatives within a market, leaving only offer characteristics (like dispatch speed) to matter in the prominence choice. However, we observe essentially all offer characteristics that the consumer observes, including dispatch speed, seller rating, whether an offer is fulfilled by Amazon (FBA), whether an offer is sold by Amazon, and the offer's price. These characteristics are accounted for in x_{jt} and do not need to be captured by the residual ξ^r_{jt}. Assumption 6 hence mainly rules out consumers using the seller's name to draw inferences about seller quality (e.g., based on prior experiences with the seller) not predicted by observables. Such a violation seems unlikely except in the case of Amazon (the merchant), which we control for by including an Amazon dummy. The remaining third-party sellers on this marketplace are numerous and often less established, and Amazon's interface makes it hard to directly search for products sold by a particular seller.

We estimate the prominence parameters $\theta^r = (\beta^r, \alpha^r, \lambda^r)$ via standard maximum-likelihood; the estimator $\hat{\theta}^r$ solves

$$\max_{\theta^r} l^r(\theta^r) = \sum_{t=1}^{T} \sum_{j \in \mathcal{J}_t \cup \{0\}} y^r_{jt} \ln r_{jt}(\theta^r), \tag{8}$$

where y^r_{jt} is an indicator for whether offer j was prominent in market t and $r_{jt}(\theta^r)$ is the model-predicted prominence probability from equation (5) but with $\xi^r_{jt} \equiv 0$.

We use maximum likelihood instead of generalized method of moments for two reasons. First, our data is high-frequency and highly disaggregated; Berry's (1994) inversion would suffer severe zero-share bias (Gandhi, Lu, and Shi 2023; Quan and Williams 2018). Second, Amazon Marketplace is highly dynamic, with frequent price changes and merchant turnover, so the time dimension provides rich, plausibly exogenous variation – especially given the frequent delegation of pricing to algorithms reacting to competitors' prices. Aggregation, the usual strategy to address zero-share bias, would thus smooth precisely over the most informative variation.

We briefly discuss identification, which is standard (Train 2009) and also confirmed by Monte Carlo exercises in Appendix D.1. The assumption that $\xi^r_{jt} \equiv 0$ implies that characteristics are exogenous (i.e., $\mathbb{E}[\xi^r_{jt}|x_{jt}] = \mathbb{E}[\xi^r_{jt}]$). Hence, the weights β^r that the algorithm places on various offer characteristics are identified by the frequency at which offers with these characteristics are featured. Furthermore, the same assumption implies that prices are exogenous (i.e., $\mathbb{E}[\xi^r_{jt}|p_{jt}] = \mathbb{E}[\xi^r_{jt}]$), and hence the price sensitivity α^r is identified by the covariance between an offer's price and prominence. Finally, it implies that the covariance between unobserved quality across offers within the "inside options" nest is zero (i.e., $\mathrm{Cov}(\xi^r_{jt}, \xi^r_{kt}) = 0$ for $k \neq j$), and hence the nesting coefficient λ^r is identified by variation in choice sets—e.g., variation in the offer number.

5.1.2 Robustness: Time-Invariant Unobserved Quality

To address concerns that price may be endogenous with respect to unobserved quality (i.e., that $\mathbb{E}[\xi^r_{jt}|p_{jt}] \neq \mathbb{E}[\xi^r_{jt}]$), we exploit the panel structure of our data. In estimation, a market t will be a product-webpage w at a particular time τ, i.e. $t \equiv (w, \tau)$. Hence, we can re-estimate the prominence model under:

Assumption 7. *Unobserved prominence quality is time-invariant:* $\xi^r_{jw\tau} \equiv \xi_{jw}$.

This alternative assumption is plausible as essentially all variation in offer characteristics is cross-sectional rather than temporal.[11] Furthermore, it explicitly allows for some endogeneity in prices (i.e., $\mathbb{E}[\xi^r_{jw}|p_{jt}] \neq \mathbb{E}[\xi^r_{jw}]$). To operationalize the assumption, we employ offer fixed-effects to concentrate out any time-invariant sources of price endogeneity.[12] However, we find below that doing so barely moves the estimate of price sensitivity, suggesting that – as expected – unobserved quality is only a minor concern in our empirical setting.

[11]The R^2 from regressing time until dispatch, log feedback count, being fulfilled by Amazon, and being sold by Amazon on offer fixed-effects are 0.96, 1.00, 0.98, and 1.00 respectively.

[12]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.

5.2 Consumer Choice

5.2.1 Main Model: No Unobserved Quality

We make the identifying assumption of no unobserved demand quality, i.e.,

Assumption 8. *There is no unobserved demand quality, i.e.,* $\xi_{jt} \equiv 0$.

As for prominence, this assumption reflects the fact that alternatives in our model are offers, and markets are product webpages; see Section 5.1.1.

We estimate the demand parameters $\theta^d = (\beta, \alpha, \lambda, \rho)$ by maximum likelihood. However, instead of observing one discrete choice per market, we observe the total number of sales $y^d_{j^o(t),t}$ for one merchant $j^o(t)$ on each market t. If we observed arrivals, we could immediately move on to the likelihood implied by our discrete choice model. As we do not observe arrivals, we utilize our assumption that arrivals are Poisson distributed, which implies that sales are also Poisson distributed. To make progress, we need to know the expected number of arrivals M_t on each market $t \equiv (w, \tau)$; to reduce noise, we assume arrival rates do not vary over time, so $M_{w,\tau} \equiv L_{w,\tau} \times \text{Arr}_w$ where $L_{w,\tau}$ measures a market's duration in months (which is observed) and Arr_w is the number of monthly arrivals. We exploit data on each market t's sales rank, translating monthly average ranks into monthly predicted sales using estimates from *AMZScout*, a "leading market intelligence provider for Amazon sellers" (Gutiérrez 2021).[13] We average across months to generate average monthly predicted sales for product webpage w. To turn predicted sales into arrivals $\hat{\text{Arr}}_w$, we use data on conversion rates of sponsored product listings, predicting each webpage's conversion rate based on its suggested retail price (higher prices convert at lower rates).[14]

Conditional on the number of arrivals $\hat{M}_t$ and the true prominence dummy y^r_{jt}, sales y_{jt} for offer j in market t follow a Poisson distribution with mean $\hat{M}_t s_{jt}(y^r_{jt}; \theta^d)$. Here, $s_{jt}(y^r_{jt}; \theta^d)$ is as specified in Equation (3) but with unobserved quality ξ_{jt} set

[13] These estimates are based on the power law relationship between sales and ranks that has been confirmed repeatedly in the academic literature (Brynjolfsson, Hu, and Smith 2006; Chevalier and Goolsbee 2003; Chevalier and Mayzlin 2006) but also account for the steepening of the slope in the tail that recent papers have found (Brynjolfsson, Hu, and Simester 2011). We are grateful to German Gutierrez for providing us with these data.

[14] The predicted conversion rates range from 10.86% (at a retail price of 99.75$) to 20.56% (at $5.29). We are grateful to Chuan Yu for providing us with these data.

to zero. We thus have a log-likelihood of

$$\sum_{t=1}^{T} \sum_{j=1}^{J_t} \left\{ y_{j,t} \ln \left(\hat{M}_t s_{j,t}(y^r_{jt}; \theta^d) \right) - \hat{M}_t s_{j,t}(y^r_{jt}; \theta^d) - \ln \Gamma(y_{j,t} + 1) \right\}, \tag{9}$$

where the last term is constant in θ^d and hence dropped below.

As written, we cannot utilize this likelihood because (i) we only observe the sales y_{jt} for one merchant $j^o(t)$ on each market t, and (ii) y^r_{jt} is only observed at the time of the last change to the set of offer characteristics $\{x_{jt}, p_{jt}\}_{j \in J_t}$ and may have changed since then.[15]

We now tackle these issues in turn. First, we note that thinning a Poisson process by classifying each arrival via a multinomial draw yields independent Poisson processes for each category (i.e. merchant j) (He and Wang 1990). Hence, we can factor out and integrate over the unobserved merchants' sales to yield an integrated likelihood

$$\sum_{t=1}^{T} \left\{ y_{j^o(t),t} \ln \left(\hat{M}_t s_{j^o(t),t}(y^r_{j^o(t),t}; \theta^d) \right) - \hat{M}_t s_{j^o(t),t}(y^r_{j^o(t),t}; \theta^d) \right\}. \tag{10}$$

Crucially, via $s_{j^o(t),t}(y^r_{j^o(t),t}; \theta^d)$, this integrated likelihood still depends on the full set of offer characteristics $\{x_{jt}\}_{j \in J_t}$ including for $j \neq j^o(t)$ – e.g., the likelihood can account for sales being lower for observed offers when they are in the presence of an attractively priced unobserved competitor by raising the coefficient on price.

Second, we address potential measurement error in the prominence dummy y^r_{jt}, which may be stale by the time of sale. Following the measurement error literature, we use an instrument that satisfies an exclusion restriction—affecting the outcome only through the mismeasured covariate (Carroll et al. 2006). In a maximum likelihood framework, this instrument informs the conditional distribution of the (true) prominence dummy, which we integrate over to obtain a full information maximum likelihood (FIML) estimator (Wooldridge 2003, Section 15.7.3).

We operationalize this idea by combining (i) the previous nested-logit prominence model, and (ii) an instrument z_{jt} for the prominence dummy that is excluded

[15] In auxiliary analyses graciously provided by Andreas Haupt, we find suggestive evidence that within fifteen minutes of the most recent change, the prominence assignment had already changed again in 52.39% of cases. Most of these high-frequency changes are unobserved by us as we only observe prominence assignments when there is a change to the offer set (e.g., a price change).

from consumer preferences: prominence eligibility (with $z_{jt} = 1$ if offer j is eligible to be featured on market t). Jointly, these yield a predicted probability $\hat{r}_{jt} = z_{jt} \times r_{jt}(\hat{\theta}^r)$ that offer j is assigned prominence on market t. We then integrate over the remaining uncertainty in prominence to yield a new log-likelihood.[16]

Putting everything together, the maximum likelihood estimator $\hat{\theta}^d$ solves

$$\begin{aligned} &\max_{\theta^d} l^d(\theta^d) = \sum_{t=1}^{T} \ln\left(\hat{r}_{j^o(t),t} \exp(\ell^d_{1,t}) + (1 - \hat{r}_{j^o(t),t}) \exp(\ell^d_{0,t})\right), \\ &\text{where } \ell^d_{r,t} = y_{j^o(t),t} \ln\left(\hat{M}_t s_{j^o(t),t}(r;\theta^d)\right) - \hat{M}_t s_{j^o(t),t}(r;\theta^d) \text{ for } r \in \{0,1\}, \end{aligned} \tag{11}$$

i.e., $\ell^d_{1,t}$ and $\ell^d_{0,t}$ are the log-likelihood contributions when the observed offer $j^o(t)$ is, respectively, prominent or not. Appendix D.3 confirms that this estimator performs well in Monte Carlo simulations.

Formally, the first-order conditions of the MLE require

$$\begin{aligned} \sum_{t=1}^{T} \Bigg[&\mathbb{P}(r_{j^o(t),t} = 1 | y_{j^o(t),t}) \left(\frac{y_{j^o(t),t}}{s_{j^o(t),t}(1;\theta^d)} - \hat{M}_t \right) \frac{\partial s_{j^o(t),t}(1;\theta^d)}{\partial \theta^d} \\ &+ \mathbb{P}(r_{j^o(t),t} = 0 | y_{j^o(t),t}) \left(\frac{y_{j^o(t),t}}{s_{j^o(t),t}(0;\theta^d)} - \hat{M}_t \right) \frac{\partial s_{j^o(t),t}(0;\theta^d)}{\partial \theta^d} \Bigg] = 0, \end{aligned} \tag{12}$$

where $\mathbb{P}(r_{j^o(t),t} = 1 | y_{j^o(t),t})$ is the posterior probability of the observed offer being prominent given $\hat{r}_{j^o(t),t}$ (the prior) and $y_{j^o(t),t}$ (the number of units sold).[17]

Intuitively, the estimator tries to find parameters such that the weighted average prediction errors are orthogonal to the derivative of predicted shares with respect to the parameters. In practice, this means that consumer preferences over characteristics β and price α are identified from the covariance between an offer's sales and its value of that characteristic relative to its competition. For instance, if the observed offer never sells when a faster dispatch offer exists, we infer high dispatch speed sensitivity. Similarly, the preference for Amazon offers is identified

[16] As prominence is binary, its (conditional) distribution is fully pinned down by the predicted probability $\hat{r}_{jt}$. By contrast, for a continuous variable, we would require full knowledge of the conditional distribution, not just the conditional mean. Formally, this procedure yields a consistent estimator if $\hat{r}_{jt}$ is well-calibrated, i.e., $\mathbb{E}[y^r_{jt} | \hat{r}_{jt}] = \hat{r}_{jt}$. This calibration is established by Figure 7a in Appendix E. Practically, the procedure performs well in Monte Carlo simulations in Appendix D.3.

[17] Formally, $\mathbb{P}(r_{j^o(t),t} = 1 | y_{j^o(t),t}) = \frac{\hat{r}_{j^o(t),t} \exp(\ell^d_{1,t})}{\hat{r}_{j^o(t),t} \exp(\ell^d_{1,t}) + (1 - \hat{r}_{j^o(t),t}) \exp(\ell^d_{0,t})}$, and analogously for $r_{j^o(t),t} = 0$.

by the shadow that the presence of an Amazon offer casts on the sales of observed merchants: if these sales are lower whenever there is an Amazon offer, and if this pattern cannot be explained by prominence alone, then consumers must prefer Amazon offers—the underlying variation used is the same as in Table 2.

We caution that this identification strategy is not without limitations. First, if Amazon systematically enters markets with *negative* market-wide demand shocks, we will attribute the lower sales on those markets to a preference for Amazon offers. However, such counterproductive entry behavior seems unlikely[18], and in any case we control for time-invariant market-wide shifters in the main demand table. Second, one may worry that we cannot separately identify the preference for Amazon offers from the substitutability between Amazon offers and observed offers: if observed offers are much more substitutable with Amazon offers than with other offers, even a slight preference for Amazon offers will cast a large shadow on the observed merchant's sales. Conversely, strong preference with lower substitutability could cast the same shadow. However, such a difference in substitutability would be identified by differing responses of the observed offer's sales to the price of the Amazon offer versus other offers. Furthermore, it seems unlikely that Amazon offers are particularly substitutable with observed offers relative to other offers. Third, with unsophisticated consumers, attributing preference correctly (to consumers vs. the prominence algorithm) requires a consistent estimate of unsophisticated consumers' prevalence; underestimating this prevalence could misattribute prominence advantage to consumer preference.

To identify the fraction of sophisticated consumers ρ, the procedure described above effectively instruments offer prominence with prominence eligibility, which sellers must purchase with a small monthly payment.[19] The estimator thus effectively targets $Cov(y^d_{j^o(t),t}, \hat{r}_{jt})$ where $\hat{r}_{jt} = r_{jt}(\hat{\theta}^r) \times z_{jt}$ is the predicted probability of being featured from the prominence model $r_{jt}(\hat{\theta}^r)$ multiplied by a dummy for prominence eligibility z_{jt}. Instrumenting in this way avoids attenuation bias from outdated prominence assignments, and also avoids potential bias from correla-

[18] If Amazon enters positive shock markets, we instead underestimate the Amazon preference. Hence, a correction would not affect our conclusion that Amazon is not self-preferencing.

[19] To be eligible, a seller must sign up for a professional account and pay a monthly subscription fee of $39.99 (junglescout.com), then become a Featured Merchant by "hit[ting] certain targets on metrics like excellent customer ratings, competitive pricing, low order defect rate, and more... [but] Amazon won't give you specific numbers" (repricerexpress.com).

tion between ξ_{jt} and ξ^r_{jt} (Donnelly, Kanodia, and Morozov 2023; Ursu 2018): the prominence model assumes $\xi^r_{jt} \equiv 0$ and hence $\hat{r}_{jt}$ does not contain ξ^r_{jt}.

The instrument is relevant by construction (first stage $F = 8{,}367$). Exclusion could be violated if an unobserved offer characteristic valued by consumers is also required for eligibility, i.e., if $\mathbb{E}[\xi_{jt}|x_{jt}, z_{jt}] \neq \mathbb{E}[\xi_{jt}|x_{jt}]$. For instance, if merchants build a reputation for on-time delivery with customers and on-time performance matters for prominence eligibility, we may misattribute the reputation-driven increase in sales to being featured. However, recall that we observe essentially all offer characteristics that consumers observe, including seller feedback, dispatch speed, whether an offer is fulfilled by Amazon, whether an offer is sold by Amazon, and its price. Hence, as before, the only plausible channel for endogeneity goes through customers drawing inferences from seller's names as they remember a particular seller's performance (e.g. for on-time delivery); this seems highly unlikely given the setting.

Finally, as with prominence, the nesting coefficient λ is identified by the slope at which sales decline as a function of the number of offers.

5.2.2 Robustness: Time-Invariant Unobserved Quality

To address the common concern that price may be endogenous with respect to quality (i.e., that $\mathbb{E}[\xi_{jt}|p_{jt}] \neq \mathbb{E}[\xi_{jt}]$), we exploit the panel structure of our data. In estimation, a market t is defined as a product webpage w observed at a particular time τ, i.e. $t \equiv (w, \tau)$. Hence, we can re-estimate the demand model under:

Assumption 9. *Unobserved demand quality is time-invariant, i.e.,* $\xi_{jw\tau} \equiv \xi_{jw}$.

As with prominence, this assumption is plausible because observed offer characteristics are essentially time-invariant. In Appendix B.2, we show how the high-frequency nature of the data allows us to condition out fixed-effects even when only one merchant's sales are observed. 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 observed merchant's sales in periods where her price is high to her sales in periods where her price is low. However, because sales are infrequent, this comparison mostly captures substitution between the observed merchant and the outside option.

5.3 Estimating wholesale and fixed costs

To ensure there are no negative realizations of costs, we assume both fixed and marginal costs follow a lognormal distribution, i.e.,

Assumption 10 (Costs DGP)**.** *For all third-party offers,*

$$\begin{bmatrix} (c_{jt})_{j\in\mathcal{N}_t} \\ F_t \end{bmatrix} = \begin{bmatrix} (\theta_0^c + x_t'\theta_x^c)(\epsilon_{jt}^c)_{j\in\mathcal{N}_t} \\ (\theta_0^F + x_t'\theta_x^F)\epsilon_t^F \end{bmatrix}, \textit{ with } \left(\begin{bmatrix} (\epsilon_{jt}^c)_{j\in\mathcal{N}_t} \\ \epsilon_t^F \end{bmatrix} | x_t\right) \sim LogN\left(0, \begin{bmatrix} \theta_\sigma^c \mathbf{I}_{|\mathcal{N}_t|} & 0 \\ 0 & \theta_\sigma^F \end{bmatrix}\right). \tag{13}$$

For the platform's offer,

$$c_{Amz,t} = (\theta_{0,Amz}^c + x_t'\theta_{x,Amz}^c)\epsilon_{Amz,t}^c, \textit{ with } \epsilon_{Amz,t}^c \sim LogN(0, \theta_{\sigma,Amz}^c). \tag{14}$$

This parameterization explicitly allows Amazon to face marginal costs that differ from those of third-party sellers. We collect the supply parameters in $\theta^s = (\theta^c, \theta^F)$ where $\theta^c = (\theta_0^c, \theta_x^c, \theta_\sigma^c, \theta_{0,\mathrm{Amz}}^c, \theta_{x,\mathrm{Amz}}^c, \theta_{\sigma,\mathrm{Amz}}^c)$ and $\theta^F = (\theta_0^F, \theta_x^F, \theta_\sigma^F)$.

Identification of θ^c follows the usual approach (Bresnahan 1981). Together, the demand model, the assumption of Nash-Bertrand competition, and variation in prices jointly identify wholesale cost parameters. In particular, the mean price across markets provides a moment that speaks to θ_0^c; its covariance with market characteristics x_t gives us θ_x^c; and the within-market price variance tells us about θ_σ^c. Their analogs for the platform offer speak to $\theta_{0,\mathrm{Amz}}^c$, $\theta_{x,\mathrm{Amz}}^c$, and $\theta_{\sigma,\mathrm{Amz}}^c$ respectively.

Meanwhile, identification of θ^F follows conventional practice in static entry models (Berry and Reiss 2007). At the entry stage, fifteen potential entrants simultaneously decide whether to enter, each knowing only its own wholesale costs and the fixed entry cost. Variation in the expected variable profits implied by the demand model—especially as driven by market size—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.

In theory, marginal cost shocks across (merchant, market) pairs ϵ_{jt}^c may depend on fixed cost shocks for that market ϵ_t^F. Following the literature on empirical two-stage games[20], the covariance matrix in Assumption 10 imposes that, conditional

[20] See, e.g., Eizenberg (2014), Holmes (2011), Houde, Newberry, and Seim (2023), Kuehn (2018),

on observables (e.g., MSRP, number of arrivals), marginal cost shocks are mean independent of fixed cost shocks. Conditioning on observables rules out many possible sources of correlation.[21]

As pricing and entry decisions are unlikely to be made at higher frequencies—and as θ^c and θ^F do not need to be identified from high-frequency variation—we aggregate to the monthly level, defining a market t as a product webpage in a given month.

We estimate $\hat{\theta}^s$ by GMM. In particular, the estimated $\hat{\theta}^s$ sets the simulated moment vector $\hat{G}(\theta^s)$ close to zero:

$$\hat{\theta}^s = \arg\min_{\theta^s \in \Theta^s} Q(\theta^s) = \hat{G}(\theta^s)' W \hat{G}(\theta^s), \quad \hat{G}(\theta^s) = \frac{1}{T}\sum_{t=1}^{T}(y_t - \widehat{\mathbb{E} f_t(\theta^s)}). \tag{15}$$

The weight matrix W is obtained by bootstrapping a covariance matrix of the underlying moments. The moments $y_t = (y_t^F, y_t^c)$ are partitioned into "fixed cost" and "wholesale cost" moments respectively, with $\widehat{\mathbb{E} f_t(\theta^s)}$ their simulated model analogs, where the expectation $\mathbb{E}$ here is taken over simulation draws. The "fixed cost" moments y_t^F are specified as

$$y_t^F = (\bar{J}_t, Var(J_t), \overline{J_t \times A_t}, \overline{J_t \times Z_{\ln(\text{MSRP}),t}}, (\overline{\mathbb{1}\{J_t = j\}})_{j=1,2,14,15})', \tag{16}$$

which includes the mean and variance of the number of entrants across markets; the mean, across markets, of the number of entrants on a product multiplied by its market size A_t; an analogous mean but with standardized MSRP (in logs) $Z_{\ln(\text{MSRP}),t}$; and finally the fractions of markets with 1, 2, 14, and 15 entrants.

Meanwhile, the wholesale cost moments, for both third-party offers and Amazon offers, take the form

$$y_t^c = (\hat{\mu}_P, \hat{\delta}_P, \hat{\sigma}_P, \hat{\mu}_{P,\text{AMZ}}, \hat{\delta}_{P,\text{AMZ}}, \hat{\sigma}_{P,\text{AMZ}})' \tag{17}$$

Rossetti (2018), and Wollmann (2018). The exposition in Bontemps, Gualdani, and Remmy (2023) is particularly salient in our setting: "Introducing such a correlation is possible in principle, but numerically challenging, because it would break the separation between the two stages in the estimation procedure. We view the absence of correlation as a reasonable simplification."

[21] Appendix D.4 performs Monte Carlo simulations to validate that (1) our procedure recovers the true parameters if the data-generating process (DGP) is correctly specified; and that (2) if wholesale cost shocks depend on fixed cost shocks in Assumption 10 other than through observables x_t, we can still recover consistent estimates of fixed costs but those for wholesale costs may be inconsistent.

and are obtained via indirect inference. In both the data and model-simulated markets, we first obtain $\hat{\sigma}_P$ by computing the standard deviation of log prices across offers. Next, we obtain $\hat{\sigma}_{P,\text{AMZ}}$ by computing the standard deviation of the price of the Amazon offer, relative to MSRP, across markets. Finally, for $(\hat{\mu}_P, \hat{\delta}_P)$, we run regressions (at the market level) predicting mean third-party offer prices as a function of a constant and product-level covariates, i.e.,

$$\bar{p}_t = \hat{\mu}_P + x_t^{P\prime}\hat{\delta}_P + \epsilon_t^P, \tag{18}$$

and we minimize the distance between the model-implied regression coefficients and the data-implied ones. The covariates for third-party offers only include the MSRP (standardized across products). To identify $(\hat{\mu}_{P,\text{AMZ}}, \hat{\delta}_{P,\text{AMZ}})$, we run an analogous regression for first-party offers, but include the price of the market's cheapest offer (possibly Amazon's own offer) as an additional covariate. This relationship between Amazon's price and the cheapest offer's price is a useful proxy for the intensity of competition.[22]

Estimation is computationally intensive. To evaluate the objective for a given guess of θ^s, we must solve the entry game for each market. However, to solve a single entry game, we must find its associated wholesale cost cutoff c_t^*. Doing so requires computing equilibrium profits under many candidate wholesale cost values. Finally, we must evaluate many candidate parameters θ^s during our outer GMM optimization procedure. We employ several techniques to lighten the computational burden; see Appendix C for technical details.

6 Results

6.1 Prominence Algorithm

We present results in Table 3, noting that all reported coefficients on offer characteristics have been normalized by dividing through by λ (i.e., they measure preferences between inside offers). While (1) assumes a (non-nested) logit model (i.e., imposes $\lambda = 1$), (2) and (3) employ a nested logit model. Finally, (4) uses the

[22]While this moment is not strictly necessary for identification, in practice, we find that it improves power: it helps the estimator distinguish prices that are low because of low costs from prices that are low due to competition.

	(1)	(2)	(3)	(4)
Price / MSRP	-7.60 (0.162)	-15.24 (0.316)	-25.17 (0.790)	-21.18 (1.577)
Time to Ship (in Days)			-1.10 (0.090)	
100 $\times$ Log(# Feedback)			5.23 (0.545)	
FBA			2.91 (0.161)	
Amazon			1.14 (0.073)	
Inside	9.55 (0.201)	4.39 (0.128)	4.39 (0.118)	
Nesting (λ)		0.14 (0.009)	0.09 (0.006)	
Offer FE?	✗	✗	✗	✓
Impl. Elasticity	-6.87	-13.30	-20.92	n/a

Table 3: Prominence Algorithm Estimation Results.

Notes: Estimates from maximum-likelihood estimation on 0.1% sample. Reported coefficients measure the effect of price (normalized by the MSRP), an extra day of time until dispatch, a percentage increase in feedback count, an offer being fulfilled by Amazon, an offer being sold by Amazon, the mean utility of the inside option (relative to the outside option), and the amount of nesting (with $\lambda = 1$ corresponding to no nesting.) The reported elasticity is the average across all offers, and cannot be computed for (4) as we condition out FE. Standard errors are clustered at the product level.

weaker identifying Assumption 7, yielding estimates robust to arbitrary correlation of an offer's average unobserved quality with its price.

We begin by discussing the price sensitivity of prominence assignments. First, correlation in the inside goods' utilities matters: moving from logit (1) to nested logit (2) doubles the estimated price coefficient. Second, including observable quality measures in (3) further increases our estimate. Third, however, additionally including offer fixed effects in (4) does not affect our estimate much.

We emphasize a few additional findings about our preferred model (3). First, the implied price elasticity for the prominence algorithm is very high, at about -21. This finding is corroborated by our discussions with sellers on the platform.[23] Second, as $\hat{\lambda} \approx 0.09 < 1$, Amazon considers offers more substitutable with each other than with the outside option. This is expected, as the outside option yields no intermediation fees. Finally, quality also matters. FBA offers have a 12% price

[23]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 are (almost) never featured.

advantage, and Amazon's offers have an additional advantage equivalent to 39% of the FBA advantage (on top of being FBA themselves). Slow shippers are penalized: a one-day longer dispatch time carries a penalty similar to Amazon's advantage.

In the Appendix, we assess the out-of-sample fit of our preferred model (3) (Figure 7, Appendix E.1) and conclude it fits well. We also explore heterogeneity (Table 14, Appendix F.1), finding the algorithm similar across categories but potentially less price-sensitive for Office and Pet products.

6.2 Consumer Choice

Our consumer choice estimation results are in Table 4, noting that all reported coefficients on offer characteristics have been normalized by dividing through by λ (i.e., they measure preferences between inside offers).

The first five columns 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 there is little endogeneity of price with respect to these covariates, implying that correlation between unobserved quality and price is also likely small.

To further assess this hypothesis, we compare a specification with offer-level fixed effects (5) to those without (1–4); model (5) is the sole model that uses the alternative identifying Assumption 9 (instead of 8). Just as for prominence, we find that including fixed effects if anything reduces the estimated price sensitivity.[24]

The remaining columns (6)-(8) freely estimate ρ, and agree that $\hat{\rho} < 1$, i.e., a substantial fraction of unsophisticated consumers only considers the prominent offer. However, (6) uses *observed* prominence assignments y^r_{jt} in place of *predicted* $\hat{r}_{jt}$; as expected, the resulting measurement error severely attenuates the estimate of the fraction of unsophisticated consumers. When correctly accounting for the measurement error in (7), our estimate of the fraction of sophisticates falls from 71% to 5%. Intuitively, this estimate should be compared with the fraction of product-page visits in which a customer clicks through to the Offer Listing Page; while we do not observe that statistic, the European Commission states that "the

[24] Amongst the non-nested models, (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).

	No Unsophisticates					With Unsophisticates		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Price / MSRP	-1.47	-1.50	-9.70	-12.40	-0.58	-11.89	-2.90	-2.99
	(0.118)	(0.097)	(0.535)	(0.576)	(0.195)	(0.628)	(0.397)	(0.376)
Time To Ship (in Days)		-0.37		-0.80		-0.63	-0.05	-0.07
		(0.053)		(0.070)		(0.079)	(0.161)	(0.152)
100 $\times$ Log(# Feedback)		8.66		-5.63		-5.68	-0.67	-0.83
		(1.330)		(1.617)		(1.875)	(3.219)	(3.116)
FBA		0.30		1.50		1.38	1.39	1.37
		(0.119)		(0.158)		(0.179)	(0.247)	(0.230)
Amazon		5.27		2.34		2.65	1.56	1.66
		(0.159)		(0.194)		(0.218)	(0.267)	(0.270)
Inside	-4.15	-4.74	-2.99	-2.97		-2.97	-3.28	-3.30
	(0.134)	(0.207)	(0.089)	(0.082)		(0.081)	(0.070)	(0.076)
$\times$ Ever Has Amz								2.22
								(2.709)
Nesting (λ)			0.11	0.10		0.09	0.04	0.04
			(0.013)	(0.009)		(0.010)	(0.041)	(0.038)
Frac. Soph. (ρ)						0.71	0.05	0.05
						(0.009)	(0.003)	(0.003)
Offer FE?	✗	✗	✗	✗	✓	✗	✗	✗
Prom. Elig. Instrument?						✗	✓	✓
Elasticity (Fix. Prom.)	-1.56	-1.60	-8.68	-10.63	n/a	-9.55	-0.98	-1.01
Elasticity (Endog. Prom.)						-11.99	-13.70	-13.72

Table 4: Demand Estimation Results.

Notes: Estimates from maximum-likelihood estimation; (6) uses observed prominence dummies y^r_{jt} in place of predicted $\hat{r}_{jt}$; (7) and (8) use predicted $\hat{r}_{jt}$ and hence account for measurement error in prominence. Reported coefficients have already been normalized by dividing through by λ (i.e., they measure preferences between inside offers.) They measure the effect of price (normalized by the MSRP), an extra day until dispatch, a percentage increase in feedback amount, an offer being fulfilled by Amazon, the mean utility of the inside option, and the change in that mean utility for markets which Amazon ever is present on. λ is the nesting coefficient for sophisticated consumers and ρ is the fraction of consumers that are sophisticated. Elasticity Fix. Prom. refers to the average price elasticity of demand assuming that price does not influence consideration sets; Elasticity Endog. Prom. allows for this influence. All standard errors are clustered at the product level.

percentage of total customer visits to the page where the customer clicked on the link to the Offer Listing Page remained below [0-10]% in the period 2017-2019" (European Commission 2022, p.10), giving us confidence in our estimate from (7).

Finally, specification (8) allows the average attractiveness of inside options to differ depending on whether an Amazon offer is ever present on a product. This is a robustness check to ensure our estimated preference for Amazon is not driven by market-level unobservables. This is the case: if anything, the preference for Amazon offers strengthens in this robustness check.

Our preferred specification (7) incorporates our consideration set model (leaving ρ unrestricted) and observable quality covariates. The price elasticity is -14,

	OLS/2SLS	MLE
Naive Estimation	51.18% (1.69%)	41.7% (0.53%)
Buybox Eligibility Instrument	90.74% (6.05%)	96.09% (0.22%)
FTC Allegation	≈98%	

Table 5: Fraction of Sales via Buybox.

Notes: This table reports the implied fraction of sales via the Buybox. Column 1 estimates this using linear regressions of sales on Buybox status, with and without an eligibility instrument. Column 2 uses our discrete choice model – after estimation, we compare predicted choice probabilities when exogenously varying Buybox status. The fraction is calculated by taking the ratio of predicted sales when in the Buybox to average sales, multiplied by the fraction of time the Buybox is occupied.

but drops to just -1 if prominence is held fixed as prices vary (i.e., if $\partial r_{jt}/\partial p_{jt}$ is artificially set to zero). This large difference stems from our estimate that 95% of customers consider only the prominent offer. However, customers considering all offers view them as close substitutes ($\hat{\lambda} \approx 0.04$).[25] Regarding non-price characteristics, we seem underpowered to detect quality effects beyond FBA and Amazon offers themselves, which are roughly equal: being sold by Amazon is worth 1.12 times as much as being FBA (intriguingly, for prominence this ratio was only 0.39). Again, note that Amazon offers are also FBA, so both advantages accrue to them.

In the appendix, we assess the out-of-sample fit of our preferred model (Figure 8, Appendix E.1), concluding that it fits well. We also explore heterogeneity (Table 14, Appendix F.1), finding that markets with more offers have more sophisticated consumers, and that the Amazon advantage is strongest for products in the Pet, Baby, and Office categories.

6.2.1 Demand Estimates Validate Regulatory Findings

The European Commission notes that (in France and Germany) "sales made directly through the Featured Offer accounted for [70-80] to [80-90] % of the

[25] If the price associated with the outside option (including competing products on Amazon) covaries with offer prices beyond what is captured by the MSRP, this estimate may be biased downwards, which may lead us to under-state the importance of entry below. However, as the inside alternatives are all offers for the same product, it is natural to expect a lot of correlation in the utility of the inside options.

overall sales" (European Commission 2022, p.10).[26] For the US, the FTC claims that "[n]early 98% of all purchases on Amazon are made using [...] the Buy Box" (Federal Trade Commission and State Attorneys General 2023, p.29).

We validate our estimates against these claims in Table 5. To this end, we estimate the causal effect of prominence on sales in two ways. First, we regress sales on Buybox status—both directly ("Naive Estimation") and using Buybox eligibility as an instrument. Each regression yields a reduced-form estimate of the effect of Buybox status on sales, allowing us to compute

$$\text{Fraction of Sales via Buybox} = \frac{\widehat{\mathbb{E}}(\text{sales}|\text{Buybox}) \times \widehat{\Pr}(\text{Buybox})}{\widehat{\mathbb{E}}(\text{sales})} \tag{19}$$

to obtain the values in Column 1 of Table 5.

Second, we use our discrete choice model of consumer behavior, comparing model-predicted consumer choice probabilities when an offer is prominent to those when it is not. Averaging over offers, this yields two structural estimates of the effect of prominence on sales (one instrumented, one not), again allowing us to compute (19), but now using the predicted choice probabilities from our model.

Omitting the instrument severely attenuates the share of sales that go through the Buybox. However, with our eligibility instrument, even the linear model recovers a sales fraction of 91%. By imposing the structure of our model, we estimate that 96% of sales on Amazon "go through" the Buybox – very close to the FTC's figure of "[n]early 98%" (ibid.).

6.3 Marginal and Fixed Costs

We present fixed and marginal cost estimates in Table 6. For successful entrants, we find wholesale costs of 81% of the manufacturer's suggested retail price (MSRP), which seems plausible for small retailers. Crucially, our estimates suggest Amazon faces significantly lower marginal costs than third-party merchants.

We estimate an interquartile range for fixed costs of [US$0.37, US$3.05]. While these fixed costs appear small, they are consistent with the low barriers to entry on Amazon. For example, 59% of merchants set prices *above* MSRP, with 15%

[26] The EC also cites Amazon's own claim that "90%+ of sales [are] com[ing] from the Buy Box" (European Commission 2022, p.10).

Parameter	Estimate	Std. Err.	Description
θ_0^F	4.676	(0.026)	mean of log fixed costs
θ_R^F	2.047	(0.049)	.. interacted with MSRP
θ_σ^F	-0.043	(0.060)	variance of log fixed cost
θ_0^c	2.863	(0.044)	marginal cost mean shift
θ_R^c	0.734	(0.006)	marginal cost coefficient on MSRP
θ_σ^c	0.154	(0.002)	variance of marginal cost
$\theta_{0,\mathrm{Amz}}^c$	3.743	(0.056)	Amazon: marginal cost mean shift
$\theta_{R,\mathrm{Amz}}^c$	0.457	(0.012)	Amazon: marginal cost coefficient on MSRP
$\theta_{\sigma,\mathrm{Amz}}^c$	0.109	(0.007)	Amazon: variance of marginal cost

Table 6: Fixed and Marginal Cost Parameters.

Notes: This table displays wholesale and fixed cost parameter estimates obtained from our SMM procedure. The estimated parameters imply an interquartile range for fixed costs of [US\$0.37, US\$3.05]. For the product at the median MSRP (\$21), the marginal cost estimates imply a mean wholesale cost of \$18.14 for potential entrants, but only wholesale costs of approximately \$16.98 for *successful* entrants. All standard errors are clustered at the product level.

pricing at higher than 1.5 times MSRP. Since the prominence algorithm 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, the model settles on low fixed costs: any larger fixed cost would render entry unprofitable for merchants with such high marginal costs.

That said, given the empirical sales frequency, our estimated fixed costs seem reasonable. These values are comparable to storage costs on Amazon. We observe roughly one sale per month for the median offer in our sample. Assuming the median merchant is storing a single item (per product), and that this item is less than one cubic foot in volume, storing it in an Amazon fulfillment center would cost about \$0.69 per product-month (junglescout.com).

7 Counterfactuals

With estimates in hand, we investigate three questions. First, we examine the value of improving search guidance. This exercise builds intuition for how the

[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.

prominence algorithm affects outcomes via its impact on the allocation of offers to unsophisticated consumers, prices, and entry. Second, this naturally leads to the question of how price-sensitive prominence should be, especially given that the platform may have an incentive to prevent sellers from competing with each other to keep intermediation fees high. Finally, we examine if steering towards Amazon's offers constitutes self-preferencing in the sense of harming consumers.

All counterfactuals assume that all markets feature a prominent offer,[28] and refer to an estimation sample of 47,249 products, representing approximately US$161 million in revenue and 465 million arrivals per month.[29] We distinguish two time horizons in our counterfactuals. In the short run, we disallow sellers from adjusting prices or their entry decisions. In the long run, sellers may change both their prices and entry decisions.

7.1 What is the value of search guidance?

We evaluate the impact of search guidance by comparing outcomes under the estimated prominence algorithm to those in a counterfactual where unsophisticated consumers only consider the cheapest offer.[30] While we do not claim this fully reflects consumer behavior in the absence of search guidance, it provides a useful benchmark for illustrating how algorithmic prominence influences market outcomes such as pricing, entry, and welfare.

As shown in Table 7, search guidance generates notable consumer gains even in the short run: relative to a scenario where unsophisticated consumers focus solely on price, the algorithm increases consumer surplus by roughly \$10 million (1.7% of Status Quo consumer surplus). This gain reflects more efficient matching: unsophisticated consumers are steered toward offers with favorable non-price attributes, such as FBA offers, that are otherwise hard to discover. In addition to raising consumer surplus, search guidance increases average third-party sales by

[28] While we can allow for cancelled prominence, doing so leads to other welfare results being overwhelmed by the negative effects of showing an unsophisticated consumer zero offers.

[29] 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.

[30] To ensure best-response iteration convergence, we assume prominence responds exclusively to price but at a price sensitivity that is only 3x the estimated sensitivity.

	Short-Run	Long-Run
Prices & Entry Adjust?	✗	✓
Δ Consumer Surplus (CS)	\$10,452,261	\$25,176,547
Δ Consumer Surplus (Naive)	\$10,452,261	\$24,390,350
Δ Consumer Surplus (Soph.)	\$0	\$786,196
Δ Producer Surplus (PS)	\$644,667	\$2,083,019
Δ Intermediation Fees	\$661,615	\$118,627
Δ Welfare (CS + PS)	\$11,758,543	\$27,378,193
Δ 3P Mean # Sales/Month	0.88	2.07
Δ 3P Mean Min Price (% MSRP)		-2.94%
Δ 3P Mean % Entrants		3.69%

Table 7: The Value of Status Quo Search Guidance (vs Seeking Lower Prices).

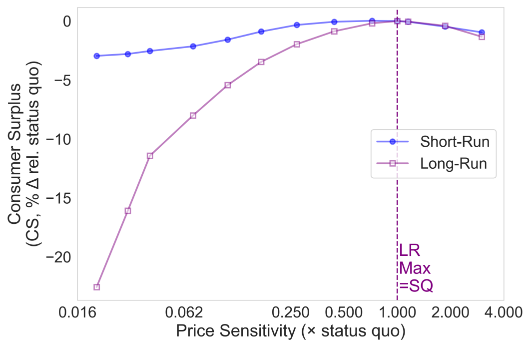
Notes: This table provides the difference in various outcomes, on our sample of 47,249 products, that can be attributed to the estimated prominence algorithm's performance relative to a non-prominence baseline where unsophisticated consumers seek lower prices. This baseline is operationalized by tripling the price coefficient in the algorithm ($\alpha^{r\prime} = 3\hat{\alpha}^r$) and switching off the dependence of the algorithm on observable non-price characteristics. In the short run, sellers cannot change their prices or entry decisions. In the long run, sellers can change both their prices and their entry decisions. Outcomes that cannot change (e.g., the number of entrants in the short run) are omitted for clarity. 3P refers to 'Third-Party', i.e., it indicates that the outcome is computed using only non-Amazon offers.

nearly one unit per product per month, thus also raising producer surplus and intermediation fee revenue.

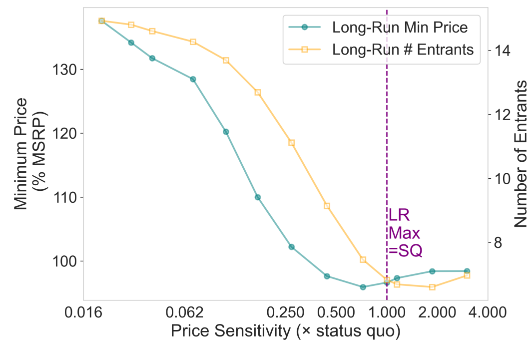
Once prices and entry are allowed to adjust, the reduced emphasis on price in the Status Quo algorithm increases producer surplus. As a result, entry increases and prices fall. Though these effects are modest, they both contribute to an additional consumer surplus increase, raising the total gains from search guidance to \$25 million (2.7% of Status Quo consumer surplus). Producers also benefit, but intermediation fees fall as a result of the decrease in prices. Finally, average monthly sales increase by 2.07 units per product.

7.2 How price-sensitive should prominence be?

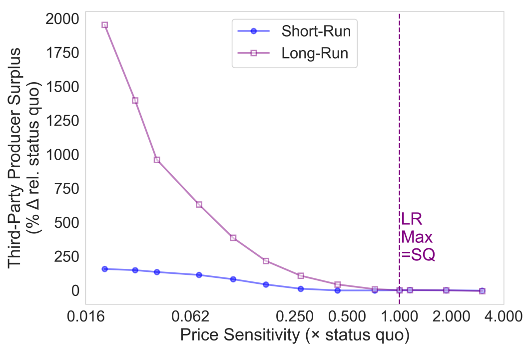
We found above that the prominence algorithm is highly price-sensitive, with a price elasticity of -21. Furthermore, when comparing the Status Quo to a world in which unsophisticated consumers only consider the cheapest offer, we found that price sensitivity plays a crucial role in redistributing surplus between producers and consumers through its impact on prices. This effect on prices flows through to Amazon's revenue via *ad valorem* intermediation fees, suggesting that Amazon may be incentivized to choose an insufficiently price-sensitive prominence algorithm.



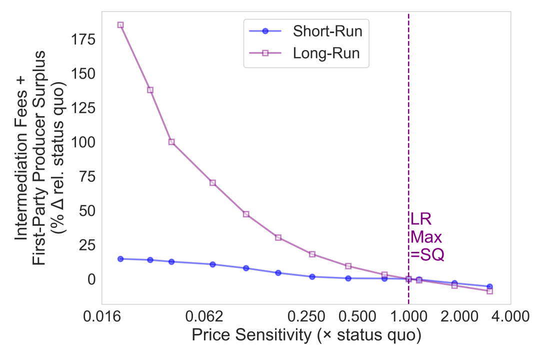
(a) Consumer Surplus



(b) Price and Number of Entrants



(c) Third-Party Producer Surplus



(d) Amazon Fees + First-Party Profit

Figure 4: How Price-Sensitive Should the Prominence Algorithm Be?

Notes: The top left panel shows the impact on consumer surplus as we vary the weight that Amazon's prominence algorithm places on price, expressed as a multiple of the Status Quo estimate $\hat{\alpha}^r$. We measure the consumer surplus impact as the % change in consumer surplus relative to the Status Quo. The top right panel shows the corresponding effects on prices and the average number of entrants. The bottom left and right panels show corresponding impacts on third-party producer surplus and Amazon profits respectively. The vertical dashed line labelled 'LR Max = SQ' represents the multiple of $\hat{\alpha}^r$ that maximizes the long-run consumer surplus, which coincides with the Status Quo price sensitivity.

However, as competition becomes more intense, participation on the platform becomes less attractive for merchants. This tradeoff raises the natural question: how price-sensitive should prominence assignments be?

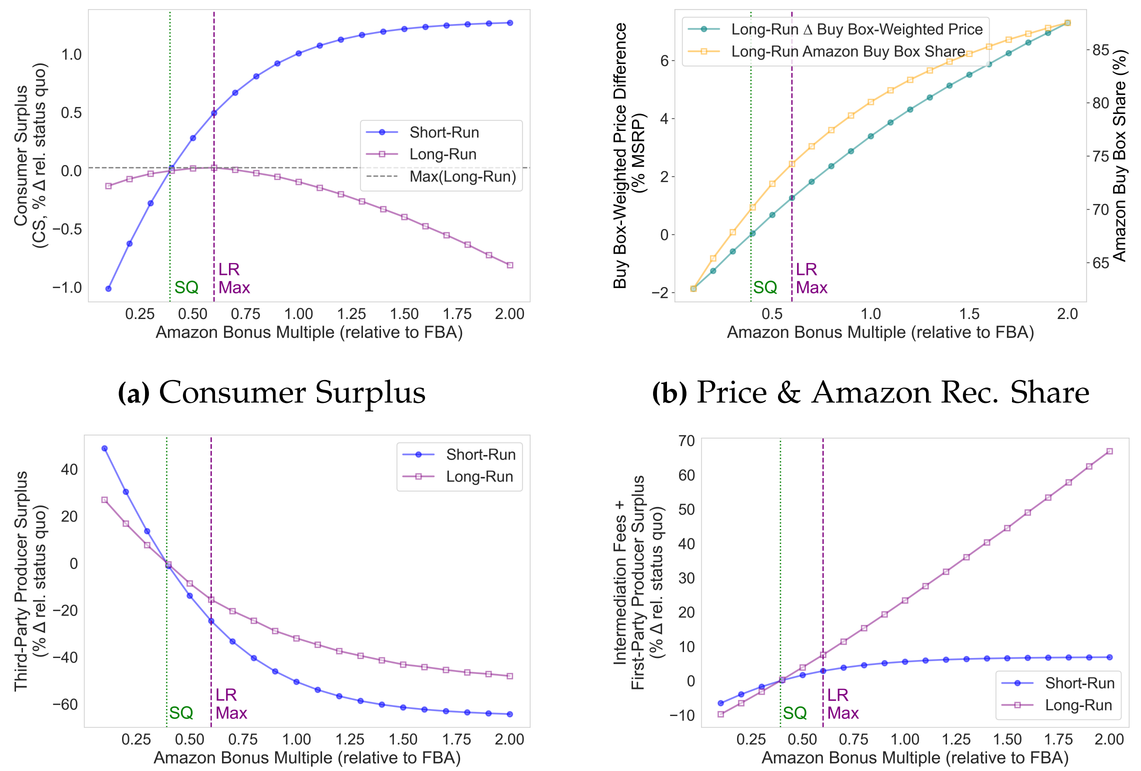
We now consider how a social planner would choose prominence assignment price sensitivity to maximize consumer surplus, accounting for endogenous pricing and entry. To do so, we vary price sensitivity (measured as a multiple of the Status Quo sensitivity) and illustrate the impacts on consumer surplus in Figure 4a; the corresponding impacts on prices and entry in Figure 4b; the effect on producer surplus in Figure 4c; and the effect on Amazon's profits (revenue from intermediation fees plus profits from first-party sales) in Figure 4d.

As Figure 4a shows, the current price sensitivity is approximately optimal from a long-run consumer surplus perspective. Our results thus do not support the view that Amazon's algorithm is inefficiently price-insensitive. Indeed, further raising price sensitivity reduces consumer surplus in both the short and long run.

How can higher price sensitivity reduce consumer surplus? First, offers on online marketplaces differ greatly in quality and service dimensions beyond price. Therefore, effectively matching the roughly 95% of unsophisticated consumers to attractive options requires the algorithm to weigh factors other than price. This "less-than-infinite" price sensitivity enables, for example, featuring Amazon-fulfilled offers despite slightly higher prices, thus benefiting consumers through better overall matches. This mechanism drives the short-run decline in consumer surplus as price sensitivity grows beyond the Status Quo (Figure 4a).

Second, driven by the decline in seller surplus illustrated in Figure 4c, an increase in the price sensitivity of prominence also decreases the number of entrants, as illustrated in Figure 4b. This effect is more pronounced at levels of price sensitivity below the Status Quo, where it mostly reduces the available choices. At higher levels of price sensitivity, it can also make competition less intense, reducing consumer surplus gains from price decreases that would result if entry had been held fixed. This second channel is responsible for the small additional decline in the long run consumer surplus (relative to the short run) as we move rightwards from the Status Quo in Figure 4a.

While increasing price sensitivity slightly reduces consumer surplus, decreasing price sensitivity does so much more dramatically, especially in the long run, where it significantly raises prices. This is concerning because it provides a mechanism through which Amazon could, in principle, exploit its intermediation power to increase profits at consumers' expense. As on-platform offers are much more substitutable with each other than with off-platform offers, prominence-driven price increases would not immediately prompt consumer switching. This "loyalty" creates room for Amazon to significantly increase revenue from intermediation fees at the expense of consumer surplus. Indeed, Figure 4d shows that decreasing price sensitivity would dramatically increase Amazon's profits, especially in the long-run as prices adjust upwards. However, our estimates indicate that, during our sample period, Amazon did not choose to increase its profits by picking such a low weight on price. Furthermore, if Amazon were to regularly feature uncompetitive



(a) Consumer Surplus **(b)** Price & Amazon Rec. Share

(c) Third-Party Producer Surplus **(d)** Amazon Fees + First-Party Profit

Figure 5: How Much Should Amazon Guide Consumers to Its Own Offers?

Notes: The top left panel shows the impact on consumer surplus as we vary the weight that Amazon's prominence algorithm places on Amazon's own offers. We measure the consumer surplus impact as the % change in consumer surplus relative to the Status Quo. The top right panel shows the corresponding effects on prices and the average share of the Amazon offer in prominence assignments, conditional on the Amazon offer being present. The bottom left and right panels show corresponding impacts on third-party producer surplus and Amazon profits respectively. The vertical dashed line labelled 'LR Max' represents the value of $\hat{\beta}^r_{Amazon}/\hat{\beta}^r_{FBA}$ that maximizes the long-run consumer surplus, while the vertical dotted line labelled 'SQ' represents the Status Quo.

offers, consumers' willingness to switch to other platforms could eventually rise (e.g., through reduced Prime membership). Such shifts in consumers' substitution patterns fall beyond the scope of our analysis.

7.3 Is Amazon self-preferencing?

Our prominence algorithm estimates point to the algorithm steering consumers toward Amazon's own offers. However, we also found that consumers prefer Amazon's offers. Therefore, without further analysis, it is unclear whether Amazon's behavior constitutes 'self-preferencing' in the sense of harming consumers.

To investigate, we run counterfactuals that vary the weight that Amazon's prominence algorithm places on its own offers.[31] We report our results in Figure 5. Figure 5a shows the impact on consumer surplus, while Figure 5b shows the corresponding effects on Amazon's prominence assignment share and the number of entrants. Figures 5c and 5d show the effects on producer surplus and Amazon's profits, respectively. The span of weights on Amazon we consider ranges from those implying an Amazon prominence assignment share from a minimum of 63% up to 88%.[32] We indicate both the Status Quo – in which the additional prominence weight on Amazon offers is 0.39 times that of FBA offers – and the value of $\hat{\beta}^r_{Amazon}/\hat{\beta}^r_{FBA}$ that maximizes long-run consumer surplus.

As expected, we find that third-party sellers unambiguously suffer from Amazon's prominence advantage – as illustrated in Figure 5c, their preferred prominence weight on Amazon offers is zero.

By contrast, the short-run consumer surplus (the blue line on Figure 5a) is locally increasing in the prominence weight on Amazon offers near the status quo (marked by the vertical line labelled 'Status Quo'): as we estimate strong consumer preferences for Amazon offers, featuring these offers more frequently is beneficial. Beyond the plotted span, this effect eventually levels off and reverses as excessive steering toward Amazon offers diverts consumers away from cheaper offers. This reversal, however, does not occur until the prominence weight on Amazon offers reaches about 5.6 times that of FBA offers. Thus, had we ignored pricing and entry decisions and looked only at the short run, we would have concluded that Amazon's current prominence weight on its offers is much too low.

In the long run, however, this is not the case. As we see in the left panel, consumer surplus is maximized at a prominence weight on Amazon offers of about 0.6 times that of FBA offers, just slightly above the Status Quo. For even higher values of $\beta^r_{Amazon}/\beta^r_{FBA}$, consumer surplus is decreasing in the prominence weight on Amazon offers, indicating that Amazon's current prominence weight is about right and Amazon is not self-preferencing in the sense of harming consumers.

What drives the difference between the short and long run? It is not entry, which remains largely stable (and is hence not plotted). Accordingly, our analysis does not support the common theory of harm that Amazon's prominence advantage

[31] Note we report results on products with an Amazon offer only as other markets are unaffected.

[32] Recall Amazon offers can be featured even if $\beta^r_{Amz} = 0$, e.g., due to competitive prices.

suppresses entry.

Instead, Figure 5b shows that the prominence-weighted price across offers rapidly increases as Amazon features its own offers more frequently. This change is almost exclusively driven by Amazon increasing its prices to exploit its larger market power, which happens because our model assumes that the platform's retail division maximizes short-run profits, just like other merchants. This assumption generates a strong pricing response by the platform to an increase in the prominence weight on Amazon offers. Indeed, we can also see this effect in Figure 5d, which shows that Amazon's profits increase more rapidly in the long run than the short run as we increase the prominence weight on Amazon offers.

However, this implication is in tension with Jeff Bezos' (admittedly conflicted) public statements that Amazon "[does] price elasticity studies, and every time the math tells us to raise prices[, but we do not]" (Rose 2013). Still, Appendix G.2 shows that fixing Amazon's prices at their Status Quo levels shifts the consumer-surplus-maximizing prominence weight on Amazon offers upward, leaving our conclusion—that Amazon is not self-preferencing—intact.

We make two cautioning observations. First, 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 market size estimates 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.

Second, our model estimates are based on data from 2018 to 2020. However, Amazon can change its algorithm at will. Hence, while our results suggest that Amazon was not self-preferencing at one particular point in time, this does not rule out the possibility that Amazon is self-preferencing at present or could do so in the future.

8 Conclusion

Regulators worry that platforms matching buyers and sellers may influence market outcomes by guiding consumer search through algorithmic prominence. We explore this issue by building and estimating a model of intermediation power. Our model captures algorithmic influence on consideration sets yet remains flexible enough to let the data speak to the extent of this influence. Our findings demonstrate the power of steering: on Amazon Marketplace, 95% of customers only consider prominent offers. Through their choice of search, ranking, and prominence algorithms, platforms are indeed able to influence market outcomes, guide consumers toward their own offers and choose the intensity of price competition. In short, algorithmic prominence gives platforms intermediation power.

However, the mere existence of intermediation power does not imply that platforms must exploit it to consumers' detriment. Indeed, we find that Amazon's prominence algorithm favors the platform. Yet the self-preferencing question is more nuanced: according to our estimates, consumers also prefer Amazon's offers. Indeed, our estimated model implies that Amazon is not self-preferencing in the sense of harming consumers. While featuring its own offers more prominently does lead to higher prices, it does not affect entry incentives much and consumers benefit from seeing offers they are likely to purchase. On net, we find that the frequency with which Amazon's algorithm features its own offers is approximately optimal from a consumer surplus perspective: featuring Amazon's offers more frequently would lower consumer surplus via higher prices, whereas featuring them less would steer consumers away from their preferred offers.

Our model also allows us to assess a recently prominent theory of harm—that self-preferencing acts as a barrier to entry. We find no evidence supporting this theory. Nevertheless, entry matters: by choosing the price sensitivity of prominence assignments, the platform can trade-off competition among incumbents against entry incentives. While some fear the platform is insufficiently motivated to foster competition because of its *ad valorem* intermediation fees, our results suggest that it has approximately chosen the consumer-surplus-maximizing price sensitivity.

We discuss future directions. On the demand side, richer sales could reveal *which* consumers are helped or hurt by platform steering. Future work could also 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 prominence algorithms is an open question. For instance, Amazon's prominence algorithm incentivises merchants to adopt FBA—does this incentive eventually harm consumers through higher prices or benefit them through better service? Finally, even large platforms face competitors. Although our entry margin implicitly accounts for merchants' ability to switch platforms, a richer model could explicitly allow both consumers and merchants to multi-home.

References

Anderson, Simon P, Andre De Palma, and Jacques-Francois Thisse (1992). *Discrete Choice Theory of Product Differentiation*. MIT Press.

Bailey, Joseph P (1998). "Intermediation and Electronic Markets: Aggregation and Pricing in Internet Commerce". PhD thesis.

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". *The Journal of Industrial Economics* 52.4, 463–496.

Ben-Akiva, Moshe and Bruno Boccara (1995). "Discrete choice models with latent choice sets". *International Journal of Research in Marketing* 12.1, 9–24.

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

Berry, Steven T (1994). "Estimating Discrete-Choice Models of Product Differentiation". *RAND Journal of Economics*, 242–262.

Berry, Steven T and Peter C Reiss (2007). "Empirical Models of Entry and Market Structure". *Handbook of Industrial Organization* 3, 1845–1886.

Blitzstein, J.K. and J. Hwang (2014). *Introduction to Probability*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press/Taylor & Francis Group.

Bontemps, Christian, Cristina Gualdani, and Kevin Remmy (2023). "Price competition and endogenous product choice in networks: Evidence from the us airline industry". Working Paper.

Bresnahan, Timothy F. (1981). "Departures from marginal-cost pricing in the American automobile industry: Estimates for 1977–1978". *Journal of Econometrics* 17.2, 201–227.

Brynjolfsson, Erik, Yu (Jeffrey) Hu, and Duncan Simester (2011). "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales". *Management Science* 57.8, 1373–1386.

Brynjolfsson, Erik, Yu Jeffrey Hu, and Michael D Smith (2006). "From Niches to Riches: Anatomy of the Long Tail". *Sloan Management Review* 47.4, 67–71.

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello (Feb. 2025). *Artificial Intelligence, Algorithmic Recommendations and Competition*.

Cardell, Scott (1997). "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity". *Econometric Theory*, 185–213.

Carroll, Raymond J., David Ruppert, Leonard A. Stefanski, and Ciprian M. Crainiceanu (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*. 2nd ed. Chapman and Hall/CRC.

Chamberlain, Gary (1980). “Analysis of Covariance with Qualitative Data”. *Review of Economic Studies* 47.1, 225–238.

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

Chevalier, Judith and Austan Goolsbee (2003). “Measuring prices and price competition online: Amazon.com and BarnesandNoble.com”. *Quantitative Marketing and Economics* 1, 203–222.

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

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

Crawford, Gregory, Matteo Courthoud, Regina Seibel, and Simon Zuzek (2022). “Amazon Entry on Amazon Marketplace”. CEPR Discussion Paper No. 17531.

Crawford, Gregory S., Robin S. Lee, Michael D. Whinston, and Ali Yurukoglu (2018). “The Welfare Effects of Vertical Integration in Multichannel Television Markets”. *Econometrica* 86.3, 891–954.

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”. *American Economic Review* 102.6, 2955–80.

Dinerstein, Michael, Liran Einav, Jonathan Levin, and Neel Sundaresan (2018). “Consumer Price Search and Platform Design in Internet Commerce”. *American Economic Review* 108.7, 1820–59.

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

Einav, Liran, Theresa Kuchler, Jonathan Levin, and Neel Sundaresan (2015). “Assessing Sale Strategies in Online Markets using Matched Listings”. *AEJ: Microeconomics* 7.2, 215–47.

Eizenberg, Alon (2014). “Upstream innovation and product variety in the US home PC market”. *Review of Economic Studies* 81.3, 1003–1045.

Ellison, Glenn and Sara Fisher Ellison (2009). “Search, Obfuscation, and Price Elasticities on the Internet”. *Econometrica* 77.2, 427–452.

Etro, Federico (2022). “The Economics of Amazon”. *Available at SSRN 4307213*.

European Commission (2022). *Commission Decision of 20.12.2022 relating to Cases AT.40462 and AT.40703*. Tech. rep. Accessed: 2025-03-31.

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

— (2025). “Vertical Integration and Consumer Choice: Evidence from a Field Experiment”. Working Paper.

Federal Trade Commission and State Attorneys General (Nov. 2023). *Complaint [Public Redacted Version]*. Tech. rep. 2:23-cv-01495-JHC. Federal Trade Commission.

Federal Trade Commission et al. (Nov. 2, 2023). *Federal Trade Commission et al. v. Amazon.com, Inc.* Complaint filed. Decision pending.

Gandhi, Amit, Zhentong Lu, and Xiaoxia Shi (2023). “Estimating Demand for Differentiated Products with Zeroes in Market Share Data”. *Quantitative Economics* 14.2, 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". *Econometrica* 76.5, 1017–1074.
Gutiérrez, Germán (2021). *The Welfare Consequences of Regulating Amazon*. Working Paper.
Hagiu, Andrei, Tat-How Teh, and Julian Wright (2022). "Should platforms be allowed to sell on their own marketplaces?" *The RAND Journal of Economics* 53.2, 297–327.
Hartzell, Olivia and Andreas Haupt (2025). "Platform Preferencing and Price Competition I: Evidence from Amazon". SSRN Working Paper No. 5126918.
He, Shengwu and Jiagang Wang (1990). "Thinning of Point Processes—Martingale Method". *Probability in the Engineering and Informational Sciences* 4.1, 117–129.
Holmes, Thomas J (2011). "The diffusion of Wal-Mart and economies of density". *Econometrica* 79.1, 253–302.
Houde, Jean-François, Peter Newberry, and Katja Seim (2023). "Nexus Tax Laws and Economies of Density in E-Commerce: A Study of Amazon's Fulfillment Center Network". *Econometrica* 91.1, 147–190.
Johnson, Justin P, Andrew Rhodes, and Matthijs Wildenbeest (2023). "Platform design when sellers use pricing algorithms". *Econometrica* 91.5, 1841–1879.
Khan, Lina M (2016). "Amazon's antitrust paradox". *Yale Law Journal* 126.
Klingler, Désirée, Jonathan Bokemeyer, Benjamin Della Rocca, and Rafael Nunes Bezerra (2020). "Amazon's Theory of Harm ". Working paper.
Kuehn, Joseph (2018). "Spillovers from entry: the impact of bank branch network expansion". *The RAND Journal of Economics* 49.4, 964–994.
Lam, Honn Tai (2021). "Platform Search Design and Market Power". Working Paper.
Luco, Fernando and Guillermo Marshall (July 2020). "The Competitive Impact of Vertical Integration by Multiproduct Firms". *American Economic Review* 110.7, 2041–64.
Malmendier, Ulrike and Young Han Lee (2011). "The Bidder's Curse". *American Economic Review* 101.2, 749–87.
Manski, Charles F. (1977). "The structure of random utility models". *Theory and Decision* 8.3, 229–254.
McFadden, Daniel (1978). "Modeling the Choice of Residential Location". In: *Spatial Interaction Theory and Planning Models*. Ed. by Anders Karlqvist, Lars Lundqvist, Folke Snickars, and Jörgen Weibull. North-Holland, pp. 75–96.
McManus, Brian, Aviv Nevo, Zachary Nolan, and Jonathan W Williams (May 2020). *Steering Incentives of Platforms: Evidence from the Telecommunications Industry*. Working Paper 27083. National Bureau of Economic Research.
Nevo, Aviv (2001). "Measuring Market Power in the Ready-to-Eat Cereal Industry". *Econometrica* 69.2, 307–342.
Quan, Thomas W and Kevin R Williams (2018). "Product Variety, Across-Market Demand Heterogeneity, and the Value of Online Retail". *RAND Journal of Economics* 49.4, 877–913.
Rasch, Georg (1960). *Studies in Mathematical Psychology: I. Probabilistic Models for Some Intelligence and Attainment Tests*. Nielsen & Lydiche.

Rasch, Georg (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". Working paper.

Reimers, Imke and Joel Waldfogel (Oct. 2023). *A Framework for Detection, Measurement, and Welfare Analysis of Platform Bias*. Working Paper 31766. National Bureau of Economic Research.

Roberts, James and Andrew Sweeting (2013). "When Should Sellers Use Auctions?" *American Economic Review* 103.5, 1830–61.

Rose, Charlie (2013). *Amazon*. Television broadcast. 60 Minutes. Aired December 1 on CBS.

Rossetti, Joseph Anthony (2018). "Product Variety in the US Yogurt Industry". PhD thesis. The Ohio State University.

Schneider, Henry S (2016). "The bidder's curse: Comment". *American Economic Review* 106.4, 1182–1194.

Smith, Michael D and Erik Brynjolfsson (2001). "Consumer Decision-Making at an Internet Shopbot: Brand Still Matters". *The Journal of Industrial Economics* 49.4, 541–558.

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

Stigler Committee on Digital Platforms (Sept. 2019). *Final Report*. Tech. rep. Stigler Center for the Study of the Economy and the State, University of Chicago Booth School of Business.

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

Teh, Tat-How and Julian Wright (2022). "Intermediation and steering: Competition in prices and commissions". *American Economic Journal: Microeconomics* 14.2, 281–321.

Train, Kenneth E (2009). *Discrete choice methods with simulation*. Cambridge University Press.

U.S. Congress (2021). *American Choice and Innovation Online Act*. H.R. 3816, 117th Congress.

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". *Marketing Science* 37.4, 530–552.

Waldfogel, Joel (Apr. 2024). *Amazon Self-preferencing in the Shadow of the Digital Markets Act*. NBER Working Paper 32299. National Bureau of Economic Research.

Wollmann, Thomas G (2018). "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles". *American Economic Review* 108.6, 1364–1406.

Wooldridge, Jeffrey M. (2003). *Econometric analysis of cross section and panel data*. Cambridge, Mass. [u.a.]: MIT Press. VII, 211.

Yu, Chuan (2024). "The Welfare Effects of Sponsored Product Advertising". Job Market Paper, Stanford University, April 2024.

Online Appendix

Two-Sided Markets Shaped by Platform-Guided Search

Kwok Hao Lee and Leon Musolff

Table of Contents

A Data Description

Our primary data source is a repricing company that administers 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 administers (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).

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) a prominence dummy, (ii) whether the offer is prominence eligible, (iii) whether Amazon fulfills it, (iv) its list price and (v) shipping price, (vi) the feedback count, (vii) the fraction of positive feedback for the associated seller, and (viii) 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 prominence* changes. Thus, while we have infinite temporal resolution on, e.g., prices, we lack this precision in prominence assignments.

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, 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. Second, we need to merge the notifications with the sales data. We do so by associating each notification with the sales we observe between the time it was sent and 48h later or the time of the next notification, whichever comes first. As this means the period during which sales are associated with a notification varies across notifications, we also record this duration for later use in market size calculations.

After merging, we obtain a dataset at the 'offer x notification'-level (with varying durations) with 204,396,069 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.[33] 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 2018-08-26 to 2020-03-23 and across 47,403 products, we observe 803,849 sales which we can attribute to a notification at most 48 hours before the sale.

	N	Avg.	Std. Dev.	Min.	50th	Max.
Price/MSRP	204,396,069	1.18	0.44	0.36	1.07	5.25
FBA	204,396,069	0.55	0.50	0.00	1.00	1.00
Amazon	204,396,069	0.02	0.13	0.00	0.00	1.00
Log (Feedback Count)	204,396,069	7.30	2.82	0.00	7.19	15.69
Time To Ship (in Days)	204,396,069	0.91	1.23	0.00	0.00	4.50
Prominence Dummy	204,396,069	0.09	0.28	0.00	0.00	1.00
Sales	19,409,013	0.04	0.57	0.00	0.00	448.00
Duration (in hours)	204,396,069	3.09	5.45	0.25	1.02	48.00

(a) Offer-Level Summary Statistics.

# Unique Products	47,403
# Merchants	63,620
# Markets (Product x Time)	19,409,013
# Sales	803,849
# Markets with Sales	411,401
Earliest Market	2018-08-26 01:18:08
Latest Market	2020-03-23 23:41:28
Earliest Market w/ Sale	2018-08-26 01:20:25
Latest Market w/ Sale	2020-03-23 22:21:50

(b) Descriptives

Table 9: Summary Statistics and Descriptives.

Notes: These tables display summary statistics for the dataset used in prominence and consumer choice estimation. For the left table, the unit of observation is an offer on a given market, so a (merchant, time, product page). 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.

B Estimation with Price Endogeneity

Our mainline estimates make use of identifying Assumptions 6 and 8 which imply $\mathbb{E}[\xi_{jt}|p_{jt}] = \mathbb{E}[\xi_{jt}]$ and $\mathbb{E}[\xi^r_{jt}|p_{jt}] = \mathbb{E}[\xi^r_{jt}]$. 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., Assumptions 7 and 9. 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 time-invariant. Indeed, the R^2 from regressing time until dispatch,

[33] This cannot influence our demand or prominence 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 prominence estimation.

log feedback count, being fulfilled by Amazon, and being sold by Amazon on offer fixed-effects are 0.96, 1.00, 0.98, and 1.00 respectively.

B.1 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 prominence algorithm estimation). Subsequently, we extend our results to the case where there is only one alternative for which we observe whether it 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 Buybox's mean utility from alternative j on product page w at time τ is given by

$$\delta^r_{jw\tau} = \mathbf{x}'_{jw\tau}\beta^r - \alpha^r p_{jw\tau} + \xi^r_{jw}, \tag{20}$$

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

Mean utilities on each offer are combined with a Type-1 extreme value shock $\epsilon_{jw\tau}$ to yield a utility index $v^r_{jw\tau} = \delta^r_{jw\tau} + \epsilon^r_{jw\tau}$. Prominence is assigned to the offer with the highest utility index. If $y^r_{jw\tau}$ is a prominence dummy, we have

$$y^r_{jw\tau} = \mathbf{1}\{v^r_{jw\tau} \geq \max_k v^r_{kw\tau}\}. \tag{21}$$

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 featured. Then,

$$\mathbb{P}\left(y^r_{jw\tau} = 1 | y^r_{jw\tau} + y^r_{j'w\tau} = 1\right) = \frac{1}{1 + \exp\left(\delta^r_{jw\tau} - \delta^r_{j'w\tau}\right)}. \tag{22}$$

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 $C = \{y^r_{jw\tau} + y^r_{jw\tau'} = 1, y^r_{jw\tau'} + y^r_{j'w\tau'} = 1\}$.

Then

$$\begin{aligned}\mathbb{P}(y^r_{jw\tau} = 1 | y^r_{jw\tau} + y^r_{j'w\tau} = 1, C) &= \frac{\mathbb{P}(y^r_{jw\tau} = 1, y^r_{jw\tau} + y^r_{j'w\tau} = 1 | C)}{\mathbb{P}(y^r_{jw\tau} + y^r_{j'w\tau} = 1 | C)} \\ &= \frac{1}{1 + \exp[(\delta^r_{jw\tau} - \delta^r_{jw\tau'}) - (\delta^r_{j'w\tau} - \delta^r_{j'w\tau'})]}.\end{aligned} \tag{23}$$

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

$$\mathcal{L}(\alpha^r, \beta^r) = \sum_{w=1}^{|\mathcal{W}|} \sum_{\tau=1}^{T_w} \sum_{j=1}^{|\mathcal{J}_{w\tau}|} \sum_{j', \tau' \in Z_{jw\tau}} \ln \left(\frac{1}{1 + \exp[(\delta^r_{jw\tau} - \delta^r_{jw\tau'}) - (\delta^r_{j'w\tau} - \delta^r_{j'w\tau'})]} \right). \tag{24}$$

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

B.2 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 $Z_{jw\tau}$ 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

$$\delta_{jw\tau} = \mathbf{x}'_{jw\tau}\beta - \alpha p_{jw\tau} + \xi_{jw}, \tag{25}$$

where, in contrast to the main text, we have imposed Assumption 9, i.e. that offer quality $\xi_{jw} \equiv \xi_{jw\tau}$ is time-invariant. Mean utilities on each offer are combined in the usual fashion with a Type-I Extreme Value shock $\epsilon_{ijw\tau}$ to yield a utility index $v_{ijw\tau} = \delta_{jw\tau} + \epsilon_{ijw\tau}$. For now, assume all consumers are sophisticated, so each consumer simply chooses her preferred option. If $y_{ijw\tau}$ is a dummy indicating whether consumer i chooses alternative j at time τ on product page w,

$$y_{ijw\tau} = \mathbf{1}\{v_{ijw\tau} \geq \max_k v_{ikw\tau}\}. \tag{26}$$

But note

$$\begin{aligned}
\mathbb{P}(y_{ijw\tau} = 1 | y_{ijw\tau} + y_{ijw\tau'} = 1) &= \frac{\mathbb{P}(y_{ijw\tau} = 1)\mathbb{P}(y_{ijw\tau'} = 0)}{\mathbb{P}(y_{ijw\tau} = 1)\mathbb{P}(y_{ijw\tau'} = 0) + \mathbb{P}(y_{ijw\tau} = 0)\mathbb{P}(y_{ijw\tau'} = 1)} \\
&= \frac{1}{1 + \frac{\sum_{k \neq j} \exp(\delta_{ikw\tau})}{\sum_{k \neq j} \exp(\delta_{ikw\tau'})} \exp(\delta_{ijw\tau'} - \delta_{ijw\tau})}
\end{aligned} \tag{27}$$

In general, $\sum_{k \neq j} \exp(\delta_{ikw\tau}) \neq \sum_{k \neq j} \exp(\delta_{ikw\tau'})$. However, when $x_{kw\tau} = x_{kw\tau'}$ and $p_{kw\tau} = p_{kw\tau'}$ 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:

$$\mathcal{L}(\alpha, \beta) = \sum_{w=1}^{|\mathcal{W}|} \sum_{(\tau, \tau') \in X_w} \ln \left(\frac{1}{1 + \exp(\delta_{ijw\tau'} - \delta_{ijw\tau})} \right), \tag{28}$$

where

$$X_w = \{(\tau, \tau') | \forall k \neq j : x_{kw\tau} = x_{kw\tau'}, p_{kw\tau} = p_{kw\tau'}\} \cap \{(\tau, \tau') | y_{ijw\tau} = 1, y_{ijw\tau'} = 0\}. \tag{29}$$

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

C Details of Simulated Method of Moments

In this section, we describe the technical details underlying the SMM procedure. 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 (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 DFO-LS, a modern, 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 (Cartis et al. 2019). This solver separately models the response of each moment to each parameter, delivering performance superior to the classic simplex algorithm (Nelder and Mead 1965).

Below, 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. Where evident, we suppress market subscripts t.

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 $q_j = \mathbf{x}_j'\beta + \xi_j$ and $q_j^r = \mathbf{x}_j'\beta^r + \xi_j^r$ measure the attractiveness of the entrant to consumers and the prominence algorithm, 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

$$\phi s_j(\mathbf{p}, \mathbf{q}) + [\phi p_j - c_j]\frac{\partial s_j}{\partial p_j} = 0. \tag{30}$$

An equilibrium of the pricing game comprises a vector of market prices $\mathbf{p}$ satisfying Equation (30) for all entrants $j \in \mathcal{J}$. To ease notation, suppress the dependence of all variables on $\mathbf{q}$, the matrix of buy box and demand qualities across merchants in a market. Following Morrow and Skerlos (2011), decompose the Jacobian matrix of market shares: $\frac{\partial \mathbf{s}}{\partial \mathbf{p}'} = \Lambda(\mathbf{p}) - \Gamma(\mathbf{p})$, where $\Lambda(\cdot)$ contains the diagonal elements of the Jacobian matrix and $\Gamma(\cdot)$ contains the factors common to both the diagonal and off-diagonal elements. Then, write the *ζ-markup equation*

$$\mathbf{p} = \phi^{-1}\mathbf{c} + \zeta(\mathbf{p}), \quad \zeta(\mathbf{p}) = \Lambda(\mathbf{p})^{-1}\Gamma(\mathbf{p})'(\phi\mathbf{p} - \mathbf{c}) - \Lambda(\mathbf{p})^{-1}\mathbf{s}(\mathbf{p}), \tag{31}$$

whenever $\Lambda(\mathbf{p})$ is nonsingular.[34]

Having described Equation (31), 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.[35]

In our model, an equilibrium exists but is unlikely to be unique. With two types of consumers, the pricing game is no longer supermodular, so the equilibrium we find depends on starting values. Our estimation procedure gives the entrant with the highest adjusted prominence algorithm attractiveness, $q_j^r - \alpha^r c_j$, a lower starting price than the others.

C.2 Solving the Entry Game

The information set[36] 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

$$\mathbb{E}_{\mathbf{q},c_{-j}}[\pi_j(c_j, q_j; \mathbf{c}_{-j}, \mathbf{q}_{-j}, \chi_{-j})] \geq F. \tag{32}$$

We focus on symmetric equilibria: $c_j^* = c^*$ for all $j \in \mathcal{N}$. This c^* satisfies a zero-profit condition at the cost cutoff:

$$V(c^*) = \mathbb{E}_{\mathbf{q},c_{-j}}[\pi_j(c^*, q_j; \mathbf{c}_{-j}, \mathbf{q}_{-j}, \chi_{-j}^*)] = F; \quad \chi_k^*(c) = 1\{c_k \leq c^*\} \; \forall k \neq j. \tag{33}$$

[34]Under some technical conditions, Proposition 2.12 in Morrow and Skerlos (2010) expresses the $\zeta(\mathbf{p})$ term as

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

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

[35]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).

[36]Henceforth, we suppress writing fixed costs F in the information set: merchants within a market t face the same fixed costs.

The LHS of this equation is a function $V(c)$ in the candidate cutoff c. We approximate $V(c)$ by an average across simulation draws, i.e.,

$$\hat{V}(c) = \frac{1}{S}\sum_{s=1}^{S} \pi_j(c^*, q_j^s; \mathbf{c}_{-j}^s, \mathbf{q}_{-j}^s, \chi_{-j}^*). \tag{34}$$

Applying standard root-finding techniques to Equation (33) yields c^*.

C.3 Simulating a Market

We combine Subsections C.1 and C.2 to simulate market-level outcomes. We employ the algorithm presented in Figure 6b. We begin by drawing a scalar fixed cost F. Next, 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 algorithm qualities $\mathbf{q}_s^r \in \mathbb{R}^{|\mathcal{N}|}$. Using the algorithm in Subsection C.2, we obtain the entry cutoff c^*. Next, referring to simulation draw $s = 1$, we find the set of successful entrants $\mathcal{J} = \{j \in \mathcal{N} : c_{js} \leq c^*\}$ and calculate their equilibrium prices, profits, and market shares using Subsection C.1.

C.4 Aggregating Markets to Moments

Following Ackerberg (2009), we aggregate market-level outcomes to moments. Fix a market t. Our model postulates a relationship f between (market-level) observables[37] $x_t = (A_t, R_t)$, unobservables u_t, and outcomes y_t. In particular, at the true parameter vector θ_0 we have

$$y_t = f(x_t, u_t, \theta_0). \tag{35}$$

We list the outcomes included in y_t in the Section 5.3 in the main text. If the data $\{x_t, y_t\}_{t=1}^T$ are generated by our model at the true θ_0, then

$$\theta = \theta_0 \implies \mathbb{E}[y_t - \mathbb{E}[f(x_t, u_t, \theta)|x_t]|x_t] = 0. \tag{36}$$

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_t. In our model, these shocks include (i) fixed

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

costs F_t for each market and (ii) qualities (q_{jt}, q^r_{jt}) as well as wholesale unit costs c_{jt} for each potential entrant on each market. To proceed, we make parametric assumptions on the distributions of these quantities. Concretely, we assume that (q_{jt}, q^r_{jt}) are drawn from the empirical distribution implied by our prominence algorithm and consumer choice estimation.

Wholesale and fixed costs are drawn following Assumption 10 in the main text. Given some candidate $\theta^s = (\theta^F_0, \theta^F_x, \theta^F_\sigma, \theta^c_0, \theta^c_x, \theta^c_\sigma)$, the conditional distribution of unobservables $p(u_t|x_t, \theta)$ is thus fully specified. In theory, we could use our model and the moment condition(s) in (36) 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 $(u_{t1}, \ldots, u_{tS}) \sim p(u_t|x_t, \theta)$ and approximate the expectation by averaging:

$$\widehat{\mathbb{E}f_t(\theta)} = \frac{1}{S} \sum_s f(x_t, u_{ts}, \theta). \tag{37}$$

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, as in Equation (15).

C.5 Concentrating Out Wholesale Cost Parameters

We partition $\theta = (\theta^F, \theta^c)$ where $\theta^F = (\theta^F_0, \theta^F_x, \theta^F_\sigma)$ and $\theta^c = (\theta^c_0, \theta^c_x, \theta^c_\sigma)$. Define

$$\tilde{\theta}^F(\theta^c) = \arg\min_{\theta_F} Q_n((\theta^F, \theta^c)), \quad \text{and} \quad \tilde{\theta}^c = \arg\min_{\theta^c} Q_n((\theta^F(\theta^c), \theta^c)). \tag{38}$$

Then, $\hat{\theta} = (\tilde{\theta}^F(\tilde{\theta}^c), \tilde{\theta}^c)$. Hence, we "concentrate out" θ^F when searching over θ^c. We use the algorithm presented in Figure 6a.

Our approach has a similar intuition to why the linear parameters are concentrated out when estimating a mixed logit model: their estimation is computationally trivial once this step is performed (Berry, Levinsohn, and Pakes 1995). For us, $\tilde{\theta}^F(\tilde{\theta}^c)$ can be found without re-solving the pricing game.

```
for p = 1,...,P do
    for s = 1,...,S_F do
        simulate entry for (p,s)
θ^F = θ^F_0
while dist > tol do
    reweight sim data for θ^F
    dist = Ĝ(θ^c, θ^F)
    update θ^F
return Ĝ(θ^c, θ^F)
```

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

```
Data: S_E sets N_{s_e} of pot entrants
for c̃ ∈ grid do
    fix c_1 = c̃
    for s_e = 1,...,S_E do
        J = {1} ∪ {j ∈ N_{s_e} : c_j < c̃, j > 1}
        find π_1(J) using fix-point iter.
    V(c̃) ≈ (1/S_E) Σ_{s_e} π_1(·)
c* solves V(c*) = F
J = {j ∈ N_1 : c_j < c*}
solve pricing game given J
```

(b) Simulation of Entry.

Figure 6: 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

To smooth over discontinuities in the entry process, we employ importance sampling Ackerberg (2009).[38] Because θ^c is fixed, we suppress mention of it in our notation. Then, write outcomes directly as a function of fixed cost draws rather than as a function of shocks: $f(x_t, u_t, \theta) = \tilde{f}(F_t)$.

We pursue the following simulation strategy: draw $F_{ts} \sim g(\cdot)$ 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

$$\widetilde{\mathbb{E}f_t}(\theta) = \sum_s \tilde{f}(F_{ts}) \frac{p(F_{ts}|x_t, \theta)}{g(F_{ts}|x_t)}. \tag{39}$$

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 $\tilde{f}(F_{ts})$ once at the beginning of the procedure. As we vary θ^F, the outcomes employed do not vary; they are simply reweighted.

[38] Finding fixed cost parameters $\tilde{\theta}^F(\theta^c)$ requires a continuous objective. However, entry is a discrete process: for instance, as we decrease F, there is only one entrant until some point at which this number "jumps" to two.

To avoid additional simulation error in our estimates, we iterate the importance sampling procedure eight times: at each new iteration, we let $g(\cdot|x_t)$ equal the density $p(\cdot|x_t, \theta^F_{prev})$ at the optimal value θ^F_{prev} of the previous iteration.[39]

D Monte Carlo Studies

To evaluate the robustness of our estimation procedure to violations of our identifying assumptions, we perform Monte Carlo analyses of demand and supply.

D.1 Monte Carlo: Prominence

Regarding prominence, the key identification concern is the possibility of correlation between unobserved prominence advantage ξ^r_{jt} and prices. Our Monte Carlo analysis is designed to evaluate the robustness of our estimator to this concern. In Table 10, we take covariates x_{jt} from our data, fix the coefficients at those estimated in our mainline analysis, and simulate the prominence outcome by combining the implied mean utilities with correlated extreme value shocks and unobserved $\xi^r_{jw\tau}$ under three assumptions:

1. Panel A: $\xi^r_{jw\tau} \equiv 0$, i.e., there is no unobserved prominence quality. Then we recover the true coefficients.

2. Panel B: $\xi^r_{jw\tau} \equiv \xi^r_{jw}$, i.e., there is unobserved prominence quality that is time-invariant but endogenous with respect to *average* prices.[40] We recover biased coefficients in the naive estimator but the offer-fixed effects estimator still recovers the true coefficients. Furthermore, unlike in our main analysis, these two estimators recover *different* coefficients, suggesting time-invariant price endogeneity is not a concern in the actual (non-simulated) data.

3. Panel C: $\xi^r_{jw\tau} = 0.4 \times p^z_{jw\tau}$, i.e., prominence quality is endogenous with respect to current prices. Then, both estimators are biased.

We conclude that time-invariant price endogeneity is not a concern (as it could be detected) but our estimators are not robust to time-varying price endogeneity.

[39] As in Ackerberg (2009), this iteration converges quickly (typically within three steps).

[40] Concretely, let $p^z_{jw\tau}$ be the z-score associated with the offer price divided by MSRP. Then if $\mathcal{T}_w$ is the set of times associated with product web-page w, we set $\xi^r_{jw} = 0.4 \times |\mathcal{T}_w|^{-1} \sum_{\tau \in \mathcal{T}_w} p^z_{jw\tau}$.

	Truth	Panel A		Panel B		Panel C	
	(0)	(1)	(2)	(3)	(4)	(5)	(6)
Price / MSRP	-24.22	-24.23 (0.049)	-22.43 (0.688)	-14.23 (0.067)	-23.65 (0.704)	-14.06 (0.029)	-13.52 (0.430)
Time to Ship (in Days)	-1.11	-1.11 (0.006)		-0.97 (0.010)		-1.12 (0.007)	
100 × Log(Feedback Count)	4.33	4.35 (0.067)		2.66 (0.115)		4.33 (0.065)	
FBA	2.56	2.54 (0.013)		2.26 (0.024)		2.53 (0.013)	
Amazon	1.11	1.11 (0.009)		0.80 (0.017)		1.11 (0.008)	
Outside	4.48	4.49 (0.012)		3.58 (0.011)		3.42 (0.009)	
λ	0.09	0.09 (0.001)		0.10 (0.001)		0.09 (0.001)	
Offer FE?		✗	✓	✗	✓	✗	✓

Table 10: Monte Carlo of Prominence Algorithm Estimation.

Notes: This table reports on a Monte Carlo exercise testing the prominence algorithm estimation. We take covariates x_{jt} from the data, fix the coefficients as in (0), and simulate the prominence outcome by combining the implied mean utilities with correlated extreme value shocks and unobserved $\xi^r_{jw\tau}$ under three assumptions. Panel A assumes that there is no demand shock, i.e., $\xi^r_{jwt} \equiv 0$. Panel B assumes that ξ^r_{jwt} is proportional to an offer's average price divided by MSRP. Panel C assumes that ξ^r_{jwt} is proportional to an offer's *current* price divided by MSRP, thus also violating the identifying assumption required for the offer-fixed effects estimator.

D.2 Monte Carlo: Demand Warmup

We now explain in a simpler model why the strategy of conditioning on a noisy proxy for the prominence dummy is valid; the notation in this section is separate from the main text. Suppose we have the following probit model

$$y_i = 1\{\alpha + \beta x_i + \epsilon_i > 0\} \tag{40}$$

where $x_i \in \{0, 1\}$ is unobserved, $\epsilon_i \sim N(0, 1)$ and (α, β) are scalar parameters. Suppose further we have a noisy proxy p_i such that $x_i | p_i, u_i \sim \text{Bernoulli}(p_i + u_i)$ where $u_i \sim F(\cdot)$ is a mean-zero unobservable. As x_i is binary, $F(\cdot)$ turns out to be irrelevant as for any $F(\cdot)$ we have $x_i | p_i \sim \text{Bernoulli}(p_i)$. Hence, we can estimate this model via maximum likelihood conditional on p_i, i.e.,

$$\begin{aligned} \ell(\alpha, \beta) = \sum_{i=1}^{n} \log \Big[& y_i \big(p_i \ell_1(\alpha, \beta) + (1 - p_i) \ell_0(\alpha, \beta) \big) \\ & + (1 - y_i) \big(p_i (1 - \ell_1(\alpha, \beta)) + (1 - p_i)(1 - \ell_0(\alpha, \beta)) \big) \Big] \end{aligned} \tag{41}$$

$$\text{where} \quad \ell_1(\alpha, \beta) = \Phi(\alpha + \beta), \quad \text{and} \quad \ell_0(\alpha, \beta) = \Phi(\alpha).$$

	$u_i \equiv 0$	$u_i \sim N(0,1)$	$u_i = \epsilon_i$
Infeasible	2.00 (0.00)	2.00 (0.00)	3.71 (0.00)
Plug-In	1.71 (0.03)	1.71 (0.03)	1.73 (0.03)
Proxy	2.00 (0.06)	1.99 (0.06)	2.03 (0.06)
True		2.0	

Table 11: Monte Carlo Simulation of Probit with Unobserved Regressor.

Notes: These Monte Carlo results compare the proposed proxy-based estimator (which integrates out the uncertainty over the unobserved regressor) with two alternative estimators: the infeasible estimator (which directly uses the unobserved binary x_i) and a plug-in estimator (where p_i is used in place of x_i in the infeasible estimator's likelihood). While the proxy estimator is unbiased, both other estimators are biased under some circumstances.

As confirmed by Monte Carlo simulations in Table 11 ('Proxy' row), this strategy results in a consistent estimator of (α, β) even when $\text{cov}(\epsilon_i, u_i) \neq 0$ as long as (i) $\text{cov}(p_i, \epsilon_i) = 0$, (ii) $\text{var}(p_i) > 0$ and (iii) p_i is well-calibrated (i.e., $\mathbb{E}[x_i|p_i] = p_i$).[41]

In our context, y_i is demand, x_i is a prominence dummy, p_i is the predicted prominence probability, ϵ_i is a demand shock and u_i is a prominence shock. Thus, even when demand and prominence shocks are correlated, we still recover the correct parameters – and that is indeed what we find in the Monte Carlo analysis below.

D.3 Monte Carlo: Demand

On the demand side, the key identification concerns are the potential endogeneity of price, as well as correlation between unobserved prominence advantage $\xi^r_{jw\tau}$ and unobserved determinants of demand $\xi_{jw\tau}$. To operationalize these concerns in a Monte Carlo analysis, we combine the recovered coefficients from our demand estimates with the covariates $x_{jw\tau}$ in our data and simulated correlated extreme value shocks to generate sales under various assumptions about demand shocks $\xi_{jw\tau}$. The results from six Monte Carlo scenarios are displayed in Table 12:

Regarding endogeneity of price:

- Panel A: No demand shock, i.e., $\xi_{jw\tau} = 0$. Then both our estimator and the one using offer fixed effects recover the true price sensitivity. Note that, as expected, the fixed effect estimator does not yield the same price coefficient

[41] Just like (i) and (ii) are reminiscent of exclusion and relevance in instrumental variables, (iii) can be thought of as a requirement on the first-stage in an IV strategy.

as reported in (0) – but this is because the estimator, which cannot distinguish between substitution between inside options and substitution to the outside option, is mixing these two sources of substitution. Indeed, as expected, the estimated price sensitivity of -1.25 in (2) lies between the *inside* price sensitivity of -3.04 reported in (1) and the implied inside-vs-outside price sensitivity in that column (which is $-3.04 \times 0.08 \approx -0.24$).

- Panel B: Price endogeneity, constant over time. Here, unobserved determinants of demand $\xi_{jw\tau}$ are proportional to an offer's *average* price divided by MSRP. Then our standard estimator is biased but the method using offer fixed effects recovers the true price sensitivity.

- Panel C: Time-varying price endogeneity. Here, $\xi_{jw\tau}$ is assumed proportional to an offer's *current* price divided by MSRP. Because identifying conditions for both estimators are violated, neither estimator recovers the true price sensitivity.

Note all of these estimators recover the correct fraction of sophisticated consumers.

Regarding prominence endogeneity & proxying:

- Panel D: There are prominence-relevant characteristics uncorrelated with price and demand shocks but which we do not observe, i.e., $\xi^r_{jw\tau} \sim N(0, 0.4^2)$. By analogy to integrating out known measurement error in a continuous variable, one may be concerned this could introduce bias for a binary variable like the prominence dummy, too.[42] However, the Monte Carlo results confirm: with a well-calibrated prominence proxy, the estimates remain unbiased (though their variance increases as the proxy becomes more imprecise). We establish in Section E, Figure 7(a) that our prominence proxy in the main analysis is well-calibrated.

- Panel E: Unobserved demand quality is correlated with unobserved prominence advantage. Here, we assume $\xi_{jw\tau} \equiv \xi^r_{jw\tau} \sim N(0, 0.4^2)$. The estimate of $\hat{\rho}$ remains unbiased, which is intuitively reasonable because the estimation

[42] As discussed in Section D.2, this concern is misleading because a binary variable's distribution is fully determined by its mean.

	Truth	Panel A		Panel B		Panel C		Panel D	Panel E	Panel F
	(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Price / MSRP	-3.00	-3.04 (0.158)	-1.25 (0.632)	7.78 (0.459)	-1.35 (0.589)	10.00 (0.243)	0.07 (0.470)	-3.48 (0.550)	-0.64 (0.624)	-5.37 (0.114)
Dispatch Time (in Days)	-0.15	-0.20 (0.064)		-0.07 (0.092)		-0.09 (0.066)		-0.00 (0.191)	0.17 (0.163)	-0.25 (0.031)
100 $\times$ Log(#Feedback+1)	-4.17	-4.26 (1.010)		-8.56 (1.669)		-5.67 (1.423)		-15.57 (5.763)	-9.76 (7.313)	0.16 (0.492)
FBA	1.23	1.06 (0.147)		0.57 (0.240)		0.65 (0.153)		2.03 (0.455)	0.49 (0.730)	0.84 (0.068)
Amazon	1.62	1.49 (0.146)		0.97 (0.820)		1.35 (0.445)		2.44 (0.609)	1.22 (0.945)	0.45 (0.082)
Outside	-3.61	-3.54 (0.020)		-4.34 (0.037)		-4.56 (0.024)		-3.60 (0.048)	-3.18 (0.167)	-3.53 (0.015)
Nesting Coefficient (λ)	0.05	0.08 (0.006)		0.06 (0.006)		0.07 (0.002)		0.04 (0.012)	0.23 (0.236)	0.05 (0.003)
Sophisticates Fraction (ρ)	0.06	0.06 (0.001)		0.06 (0.004)		0.06 (0.002)		0.06 (0.005)	0.06 (0.025)	0.26 (0.003)
Offer FE?		✗	✓	✗	✓	✗	✓	✗	✗	✗

Table 12: Monte Carlo of Demand Estimation.

Notes: This table provides results from a Monte Carlo exercise that tests our demand estimation. We sample covariates X from our real data and use the coefficients in (0) to simulate sales under various assumptions about demand shocks ξ_{jwt}. All panels assume a 75% probability that prominence assignments changed since the last observation. Panel A assumes no demand shock, i.e., $\xi_{jwt} \equiv 0$. Under this assumption, both the naive and offer fixed-effects estimators recover the true price sensitivity (i.e., coefficient on Price/MSRP). Panel B assumes that ξ_{jwt} is proportional to an offer's *average* price divided by MSRP. The naive estimator is biased, but we recover the true coefficients once we include offer fixed-effects. Panel C assumes that ξ_{jwt} is proportional to an offer's *current* price divided by MSRP, thus violating the identifying assumption required for the offer-fixed effects estimator; indeed, our estimates are now biased. Panel D assumes var$(\xi^r_{jwt}) > 0$, which does not bias estimates of ρ. Panel E assumes that ξ_{jwt} is exogenous with respect to price, but $\xi_{jwt} \equiv \xi^r_{jwt}$; this prominence endogeneity does not bias $\hat{\rho}$. Panel F is like Panel A but replaces the eligibility instrument with the short-run correlation between prominence and sales; the mismeasurement of prominence leads to an upward bias in the fraction of sophisticates.

strategy purges any endogenous variation in prominence status by projecting it only on known variables (the intuition is the same as with an instrumental variables strategy.)

- Panel F: Similar to Panel A, but we replace the eligibility 'instrument' with the short run correlation between prominence and sales. The mismeasurement of prominence leads to an upward bias in the share of sophisticated consumers.

To summarize, we find that our estimates are robust to prominence endogeneity, and would detect price endogeneity that is constant over time. However, they are biased if price endogeneity varies over time, and would be biased if we used the *observed*, rather than the *predicted*, prominence dummy (which we do not).

D.4 Monte Carlo: Supply

We also verify our supply model by Monte Carlo simulations. Our exercise achieves two goals. First, under our maintained assumptions, we can recover the correct fixed and both third- and first-party wholesale cost parameters. Second, if costs are misspecified because there is unmodelled correlation between fixed and wholesale costs, then – as expected – both fixed and wholesale cost parameter estimates are biased.

Results are displayed in Table 13. Columns marked A fix all parameters at their true values, except for the mean fixed cost parameter μ_F and the mean wholesale cost parameters $(\mu_c, \mu_{c,\text{Amz}})$. Columns marked B allow all parameters to vary. The columns (A1) and (B1) validate that we can recover the fixed and wholesale parameters we impose. In columns (A2) and (B2), we estimate the model under misspecified wholesale costs.

For exposition, consider a product t and a seller j. To simulate fixed and wholesale costs, we first draw fixed cost "shocks" $\epsilon_{F,t}$ from a standard normal. Next, we draw wholesale cost "shocks" following $\epsilon_{c,jt}|\epsilon_{F,t} \sim \Sigma\epsilon_{F,t} + (1-\Sigma)N(0,1)$, for some value of Σ in $[0,1]$. The value of Σ is set to zero for columns marked 1, and 0.5 for columns marked 2. Intuitively, a higher positive value of Σ means that wholesale cost shocks are higher on products with higher fixed costs (i.e., fewer entrants).

In columns (A1) and (B1), we find that our estimator delivers excellent point estimates when the model is correctly specified, and the associated confidence intervals are very tight around the truth. However, for the misspecified model, if we fix all parameters other than $(\mu_F, \mu_c, \mu_{c,\text{Amz}})$ at their true values, we see in column (A2) that we obtain highly biased estimates for third-party wholesale costs and somewhat biased estimates for mean fixed costs. If we permit all parameters to vary, as in column (B2), both fixed and wholesale cost parameters are now biased away from the truth.

E Out-of-Sample Fit

We display out-of-sample fit for the Buybox, demand, and supply respectively. In each case, we perform an 80-20 train-test split: we train the model on 80% of the data, then test it on the remaining 20%. On all fronts, we find that our models fit the held-out data well.

	Truth	(A) (A1)	 (A2)	Truth	(B) (B1)	 (B2)
μ_F	4.647	4.655 (0.010)	4.462 (0.007)	4.647	4.660 (0.014)	4.746 (0.016)
$\mu_{F,M}$	1.921			1.921	1.920 (0.031)	1.853 (0.033)
σ_F	-0.049			-0.049	-0.021 (0.024)	0.305 (0.014)
δ_c	0.777			0.777	0.770 (0.004)	0.780 (0.004)
σ_c	0.156			0.156	0.157 (0.001)	0.099 (0.001)
μ_c	2.258	2.267 (0.011)	4.955 (0.014)	2.258	2.395 (0.032)	2.895 (0.063)
$\delta_{c,AMZ}$	0.474			0.474	0.461 (0.007)	0.482 (0.004)
$\sigma_{c,AMZ}$	0.117			0.117	0.147 (0.017)	0.004 (0.015)
$\mu_{c,AMZ}$	3.487	3.490 (0.033)	3.454 (0.019)	3.487	3.475 (0.065)	4.614 (0.065)

Table 13: Monte Carlo of Supply Estimation.

Notes: This table provides results from a Monte Carlo exercise that tests our code which estimates fixed and wholesale cost parameters. We generate a dataset by simulating markets from 47,249 products in our real data. Using upstream demand parameters, the coefficients reported in (0), as well as simulated correlated shocks for fixed and wholesale cost parameters, we generate the dependent variable under different assumptions about the correlation between fixed and wholesale costs. First, fixed cost "shocks" $\epsilon_{F,t}$ are drawn from a standard normal, then wholesale cost "shocks" are drawn $\epsilon_{c,jt}|\epsilon_{F,t} \sim \Sigma\epsilon_{F,t} + (1-\Sigma)N(0,1)$, for some value of Σ in $[0,1]$. These shocks are then transformed into fixed costs via $F_t = (\mu_F + \mu_{F,M} z_{\text{MSRP},t}) \exp(\sigma_F \epsilon_{F,t})$, where $z_{\text{MSRP},t} = \ln \text{MSRP}_t - T^{-1} \sum_{t=1}^{T} \ln \text{MSRP}_t$. Wholesale costs are obtained via $c_{jt} = (\mu_c + \delta_c \times \text{MSRP}_t) \times \exp(\sigma_c \epsilon_{c,jt} - \sigma_c^2/2)$, and analogously for first-party offers. Panel A fixes all parameters except the displayed parameters at their true values; Panel B allows all parameters to vary. Specifications A1 and B1 use data simulated under the assumption of independent wholesale and fixed costs; the resulting estimates are close to the true parameters. Specifications A2 and B2 simulate correlated wholesale and fixed costs; as expected, the estimates that our now misspecified estimator yields in this case are biased.

E.1 Out-of-Sample Fit: Prominence and Demand

We analyze the fit of the demand and prominence models by comparing their out-of-sample predictions to the data. To this end, we first estimate the models on a sample of 80% of products. Then, for each product in the remaining 20%, we predict (i) each offer's probability of being promoted, and (ii) each observed offer's sales.[43] We then compare these predictions to the data. While we could assess fit on the moments that our maximum-likelihood estimation is implicitly targeting, we could only miss these moments if the model was overfitting; we confirm in

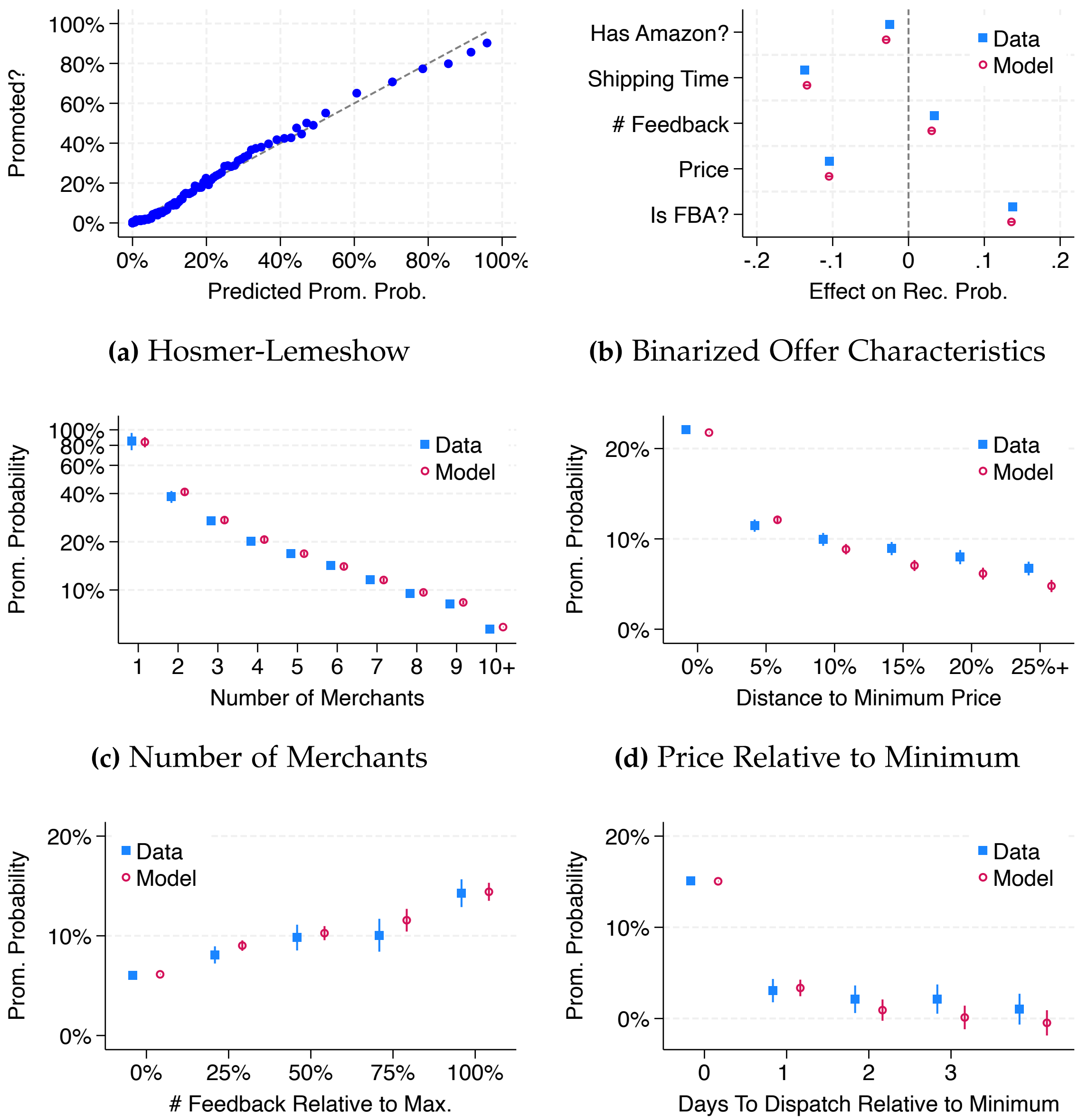
[43] Note that predicted sales are a function of prominence assignments at the time of sale, which we do not observe; we use the posterior probability of (not) being promoted given our model as the weight on the predicted sales conditional on (not) being promoted.

unreported tests that this is not the case, i.e., the fit on these moments is good. We focus here instead on more substantive measures of fit.

Our results are Figures 7 (for prominence) and 8 (for sales). Both figures first show a Hosmer-Lemeshow test that speaks to the calibration of the models, binning observations by their predicted value of the outcome variable and comparing the distribution of the outcome variable to the distribution of the predicted values. We find that these predictions lie close to the 45-degree line, indicating that both models are well-calibrated, with the possible exception of the demand model's slight tendency to under-predict sales of the lowest-selling offers. The excellent calibration of the Buybox model is particularly reassuring given that, as we discuss in Appendix D.3, the identification of the fraction of sophisticates in the demand model relies on the calibration of the Buybox model.

The second panel shows the results of regressing (without any additional controls) the predicted and actual outcomes on binary or binarized offer characteristics (with the exception of "In Buybox?" which uses predicted prominence given that the observed prominence assignment is mismeasured). We find that the models are able to capture the relationship between the outcome variables and the offer characteristics well. We note in particular that the impact of being in the presence of an Amazon offer is matched well by the models.

Finally, the remaining panels try to assess whether our models correctly capture the causal relationship between offer characteristics and prominence/sales. To this end, we regress both model predictions and data on various offer characteristics while controlling for offer fixed effects. If (i) the model is correctly specified and (ii) temporal variation in offer characteristics is uncorrelated with the error term, then the coefficient on the offer characteristics in the two regressions should be approximately equal. Furthermore, this exercise allows us to verify that both the data and the model capture intuitive patterns (e.g., that sales rapidly decline as price increases beyond the minimum price on the market but decline more slowly for further price increases). The models capture the relationship between the outcome variables and the offer characteristics well. We note in particular that, thanks to the nested-logit structure of the models, the relationship between the number of merchants on a market and the probability of being promoted is matched well by the models.



(a) Hosmer-Lemeshow

(b) Binarized Offer Characteristics

(c) Number of Merchants

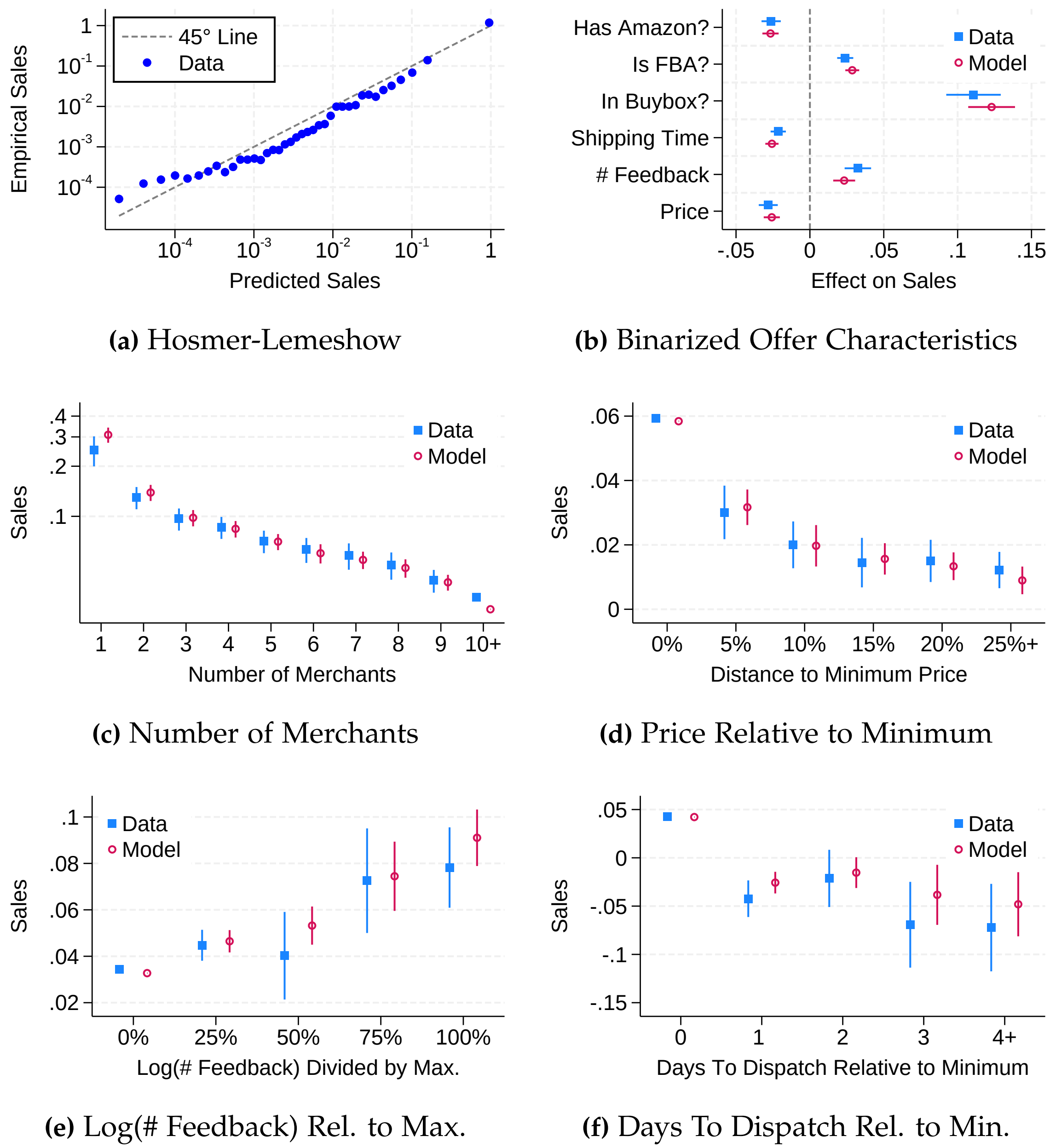
(d) Price Relative to Minimum

(e) Log(# Feedback) Rel. to Max.

(f) Days To Dispatch Rel. to Min.

Figure 7: Out of Sample Fit for Prominence Model.

Notes: This figure compares the prominence model's out-of-sample predictions to the data. Panel (a) shows a Hosmer-Lemeshow test, binning observations by predicted probability of being promoted and comparing to actual prominence observations. Panel (b) shows how well the model captures the relationship between prominence assignment (i.e., Buybox wins) and binary offer characteristics like FBA status; continuous characteristics are binarized by comparing to the median value. The underlying regression includes just the dummy and no other controls. Panels (c)-(f) show results from regressing the observed prominence dummy and predicted probabilities on various market- or observed offer characteristics while controlling for offer- fixed effects. Panel (c) examines fit across markets with different numbers of competing merchants. Panel (d) evaluates how well the model captures the relationship between an offer's relative price position and its likelihood of being promoted. Panels (e) and (f) repeat this exercise for feedback count and time to dispatch respectively. All standard errors are clustered at the product level.



(a) Hosmer-Lemeshow

(b) Binarized Offer Characteristics

(c) Number of Merchants

(d) Price Relative to Minimum

(e) Log(# Feedback) Rel. to Max.

(f) Days To Dispatch Rel. to Min.

Figure 8: Out of Sample Fit for Demand Model.

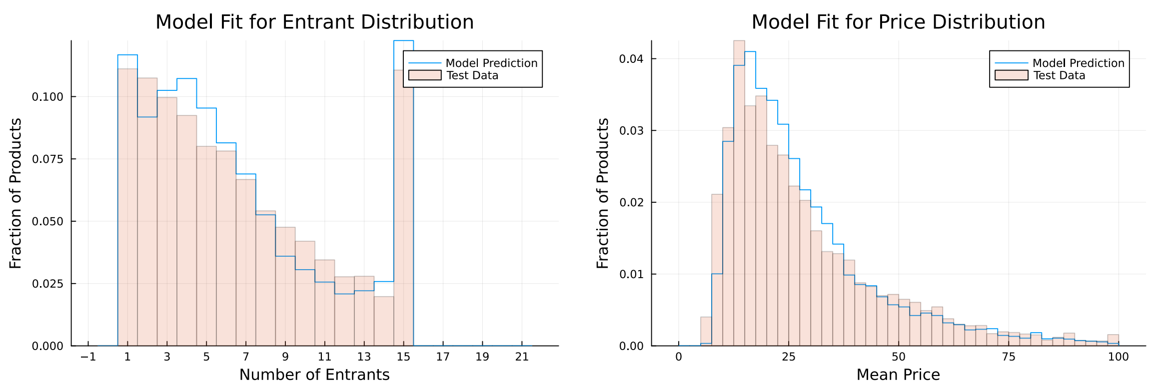
Notes: This figure compares out-of-sample predictions of our demand model to the data. Panel (a) shows a Hosmer-Lemeshow test, binning observations by predicted sales and comparing to actual sales. Panel (b) shows how well the model captures the relationship between sales and binary offer characteristics like FBA status; continuous characteristics are binarized by comparing to the median value, except for "In Buybox?" which uses the predicted prominence given that the observed prominence is mismeasured. The underlying regression includes just the dummy and no other controls. Panels (c)-(f) show results from regressing actual sales and sales predictions on various market- or observed offer characteristics while controlling for offer fixed effects. Panel (c) examines fit across markets with different numbers of competing merchants. Panel (d) evaluates how well the model captures the relationship between an offer's relative price position and its sales. Panels (e) and (f) repeat this exercise for feedback count and time to dispatch respectively. All standard errors are clustered at the product level.

E.2 Out-of-Sample Fit: Supply

We analyze the fit of the supply model by comparing its out-of-sample predictions of the price and entrant distribution to those in the data.

As for the demand model, we first estimate the supply model on a "training set" comprising 80% of the products. Then, for each product in the remaining 20%, we simulate entry and pricing under the estimated parameters. Finally, we compare the simulated number of entrants and prices to the data.

We show our results in Figure 9. We find that the model is able to replicate the empirical distribution of entrants and prices.



(a) Entrant Distribution **(b)** Price Distribution

Figure 9: Out of Sample Fit for SMM Model.

Notes: This figure compares out-of-sample predictions of our entry model to the data. Panel (a) shows the model predicted entrant distribution as an outline overlaid on the entrant distribution in the test data. Panel (b) shows the analogous price distributions for the test data. The entry model is fitted on an 80% random sample of products; this model is then used to simulate market outcomes on the remaining 20% of the data held out from the training set.

F Heterogeneity Analysis

To evaluate the heterogeneity of our results across various partitions of our data, we estimate our model separately for each product category.

F.1 Heterogeneity: Prominence and Demand

In Table 14, we investigate whether the key parameters of our demand and prominence models vary by (i) product category, or (ii) product characteristics. To this end, we re-estimate prominence and demand models separately for each subset of the data defined by a heterogeneity dimension (first column) taking on a particular value (second column).[44] Instead of reporting estimates directly, we report the following key implied parameters: Amazon's prominence advantage (as multiple of the FBA advantage) in Column (1), the price elasticity of prominence in Column (2), Amazon's price advantage in demand (as multiple of the FBA advantage) in Column (3), the price elasticity of demand in Column (4), and the percentage of sophisticated consumers in Column (5).

To stabilize the estimates, we fix the nesting parameter λ and the impact of all characteristics other than price and being sold by Amazon at their full-sample estimates as these parameters are poorly identified in smaller samples. To obtain standard errors, we perform a parametric bootstrap.

Starting with heterogeneity across product categories, we find some evidence of heterogeneity in both demand and prominence. In particular, Amazon's demand advantage is higher for products in the Pet, Baby and Office categories, and lower for products in the Fashion, Sports and Books categories. Furthermore, there is a tendency for categories with high demand advantage to have higher prominence advantages, though the relationship is noisy. Overall, the pattern (high Amazon advantages for goods needed quickly) suggests that Amazon offers arrive at the customer's doorstep faster in a way that is not captured by our dispatch time variable – perhaps because Amazon offers qualify for Prime at a higher rate, or because Amazon's offers spend less time in transit even conditional on Prime status.

Moving on to price elasticities, both prominence and demand are more elastic for Food products, and less elastic for Office products, which makes sense if Office products are mostly purchased by people spending their employer's money.

Finally, regarding the fraction of sophisticated consumers, we find that consumers are more likely to explore alternative offers especially when buying books and health or beauty products. They are least likely to explore when buying tools, sporting goods or office products. This could be driven by heterogeneity in willingness to pay for faster shipping for books (and unwillingness to spend one's time to help one's employer save money in the case of office products).

Moving on to binary product characteristics, we find that Amazon's demand

[44] The predicted prominence probabilities required to estimate our demand model are also re-estimated for each subset.

advantage is higher for products with a lower MSRP – plausible, as this advantage is implicitly expressed as a percentage of MSRP, and that percentage would be a larger absolute amount for higher-priced products. The biggest difference when it comes to the fraction of sophisticated consumers is that products with many offers attract more sophisticated consumers – and these products also feature a much elevated prominence price elasticity. Again, these findings are reasonable: there is a higher payoff from examining all offers when there are more offers to examine, and from Amazon's perspective, there is no need to worry about incentivizing entry of additional merchants when there are already many offers.

To conclude, we emphasize that this heterogeneity analysis should be interpreted with caution due to the convenience sample of products that underlies our analysis. While we have good coverage of different kinds of sellers across categories, once we zoom in on a particular category or heterogeneity dimension, we are left with only a small number of products, limiting power. In particular, the underlying estimates of the price/MSRP coefficient in demand, α, are not always statistically distinguishable from zero. Furthermore, after conditioning on a heterogeneity dimension, the remaining products are not necessarily representative of the population of products in that category or heterogeneity dimension. For both these reasons, the results in this section and the next one should be interpreted with caution.

F.2 Heterogeneity: Implications for Counterfactual Analyses

What do our findings suggest for the heterogeneous impacts of search guidance, as well as potential antitrust remedies, across product categories? First, we re-estimate the pricing and entry model in the body of the paper assuming the Buybox and demand processes in each category are governed by the parameters recovered in Panel A of Table 14.[45] Next, we repeat the main counterfactual analyses from Section 7. Finally, we display the effects of these interventions on consumer surplus in the aggregate within category, as well as per product within that category.

The results are summarized in Table 15. The pooled consumer surplus results are qualitatively similar to those of the main counterfactual: search guidance helps

[45]To stabilize SMM convergence, our indirect inference for first- and third-party wholesale prices aims to match only the constant and standardized MSRP coefficients in their respective regressions of offer prices on market characteristics.

Dimension	Value	Prominence (1)	(2)	Demand (3)	(4)	(5)
Category	Baby	0.32	-22.90	2.84	-13.43	7.89
		(0.21)	(4.83)	(0.41)	(1.83)	(3.01)
Category	Beauty	0.15	-18.61	1.91	-14.06	23.88
		(0.04)	(0.68)	(0.18)	(0.66)	(1.67)
Category	Books	0.42	-21.92	0.56	-10.18	26.57
		(0.05)	(0.99)	(0.75)	(0.77)	(6.71)
Category	Fashion	0.59	-24.44	0.99	-17.69	4.88
		(0.10)	(1.36)	(0.68)	(0.78)	(0.74)
Category	Food	0.29	-31.18	2.19	-16.56	14.99
		(0.09)	(1.36)	(0.23)	(0.84)	(2.98)
Category	Health	0.35	-21.00	1.49	-15.27	22.74
		(0.08)	(1.07)	(0.30)	(0.95)	(3.39)
Category	Home	0.17	-19.26	1.92	-13.33	6.82
		(0.04)	(1.20)	(0.21)	(0.62)	(0.95)
Category	Office	0.84	-12.51	2.45	-9.80	8.87
		(0.17)	(1.40)	(0.20)	(0.97)	(1.66)
Category	Pet	1.07	-14.11	2.95	-10.49	10.13
		(0.43)	(2.22)	(0.42)	(1.22)	(1.56)
Category	Sports	0.27	-25.31	0.81	-18.08	3.28
		(0.05)	(1.48)	(0.46)	(1.02)	(0.54)
Category	Tools	0.48	-18.23	1.35	-13.25	4.89
		(0.14)	(1.39)	(0.34)	(0.84)	(0.98)
Category	Toys	0.95	-24.29	2.51	-12.64	11.08
		(0.13)	(0.92)	(0.16)	(0.49)	(0.92)
Category	Video Games	0.45	-23.88	2.32	-12.03	10.27
		(0.53)	(2.48)	(0.47)	(0.77)	(2.19)

Panel A: Heterogeneity by Product Characteristics.

Dimension	Value	Prominence (1)	(2)	Demand (3)	(4)	(5)
Durable	No	0.33	-21.15	1.08	-14.32	5.79
		(0.04)	(0.73)	(0.16)	(0.37)	(0.37)
Durable	Yes	0.42	-21.53	1.15	-13.63	3.48
		(0.04)	(0.69)	(0.17)	(0.36)	(0.37)
High item volume	No	0.50	-19.29	1.39	-12.52	6.88
		(0.05)	(0.68)	(0.14)	(0.34)	(0.46)
High item volume	Yes	0.48	-22.82	0.96	-14.10	3.79
		(0.04)	(0.97)	(0.22)	(0.42)	(0.43)
High msrp	No	0.43	-20.19	1.51	-12.49	8.75
		(0.04)	(0.60)	(0.14)	(0.29)	(0.59)
High msrp	Yes	0.45	-23.80	0.75	-15.93	3.49
		(0.04)	(0.85)	(0.22)	(0.42)	(0.26)
High num offers	No	0.40	-15.45	1.07	-11.67	3.39
		(0.04)	(0.60)	(0.15)	(0.35)	(0.21)
High num offers	Yes	0.44	-27.64	1.84	-14.70	12.01
		(0.04)	(0.77)	(0.14)	(0.31)	(0.80)
High package weight	No	0.42	-20.94	0.68	-13.80	4.30
		(0.04)	(0.64)	(0.19)	(0.33)	(0.30)
High package weight	Yes	0.48	-20.96	1.35	-13.59	5.66
		(0.04)	(0.79)	(0.16)	(0.35)	(0.43)
Working hours	No	0.41	-20.82	1.19	-13.65	4.78
		(0.04)	(0.79)	(0.12)	(0.37)	(0.27)
Working hours	Yes	0.40	-21.81	1.09	-13.95	5.29
		(0.03)	(0.71)	(0.12)	(0.31)	(0.28)

Panel B: Heterogeneity by Product Characteristics.

Table 14: Heterogeneity of Prominence and Demand Models

Notes: We re-estimate prominence and demand models separately for each subset of the data defined by a heterogeneity dimension (first column) taking on a particular value (second column). Column (1) reports the Amazon advantage in the prominence algorithm (expressed as a multiple of the FBA advantage). Column (2) reports the average price elasticity of prominence. Column (3) reports the Amazon advantage in demand (again expressed as a multiple of the FBA advantage). Column (4) reports the price elasticity of demand (taking into account the impact of prominence responding to price). Column (5) reports the percentage of sophisticated consumers. To dichotomize products into durable and non-durable, we asked Claude Sonnet 3.5 to estimate the fraction of products sold in each category on Amazon that are durable; we consider a category durable if the fraction is above 50%. The item volume refers to the product of the width, height and length of the product. Working hours refer to 9AM to 5PM in Chicago, excluding weekends. All binarized variables are dichotomized at the median value of the full sample unless otherwise indicated. Standard errors are clustered at the product level.

consumers; softening price competition harms them; and handicapping Amazon fails to deliver substantial consumer surplus gains. Henceforth, we focus on the differential impacts by category. First, search guidance is most valuable to goods in the Baby (33.4%), Fashion (23.1%), Video Games (13.1%), and Books (10.6%) categories, though not for Office products (-1.0%). Second, making comparisons on a per-product basis, handicapping Amazon may improve consumer surplus for Office products (2.0%), but is unlikely to deliver substantial gains in any other category. Finally, if Amazon were to soften price competition to extract more seller fees, consumer surplus declines the most for products in the Office (-2.5%), Beauty (-1.7%) and Health (-1.5%) categories, but may benefit consumers of Books somewhat (2.2%).

As before, we caution against taking these results as precise estimates of welfare impacts across categories, given our limited sales data and hence noisy category-level consumer choice estimates.

Category	N	Value of Guidance	Handicap Amazon	Soften Pricing
Baby	529	33.4%	0.12%	-0.43%
Beauty	4220	2.79%	0.33%	-1.74%
Books	5676	10.64%	-0.1%	2.18%
Fashion	12766	23.09%	0.02%	0.09%
Food	4259	2.06%	0.19%	-1.11%
Health	2502	1.27%	0.61%	-1.51%
Home	3393	2.77%	0.08%	-0.81%
Office	1186	-0.96%	1.97%	-2.42%
Pet Supplies	667	1.14%	0.42%	-1.28%
Sports	3653	5.09%	0.06%	-0.39%
Tools	1075	3.5%	0.44%	-1.41%
Toys	6210	2.69%	0.43%	-0.97%
Video Games	571	13.05%	0.2%	-0.29%

Table 15: Heterogeneous Effects on Consumer Surplus (% Relative to Baseline)

Notes: We simulate counterfactuals separately for each subset of the data defined by a product category taking on a particular value (first column). The number of products in that category, N, is also displayed (second column). After using the prominence and demand estimates in Panel A of Table 14, we estimate an entry and pricing process as in the main text, then simulate three main counterfactuals: the value of search guidance relative to if consumers sorted strongly on price, i.e., if the prominence algorithm had three times the price sensitivity and did not select on non-price characteristics (Column 3); forbidding Amazon from promoting its own offer above and beyond what is predicted by our observables (Column 4); and if Amazon halved its price coefficient to soften price competition (Column 5). All consumer surplus figures reported are from the "Long-Run" scenario in the main text, where merchants adjust both pricing and entry decisions. Pooled consumer surplus results are suppressed because they are qualitatively similar to those of the main counterfactual.

G Counterfactuals

G.1 Computing Counterfactual Outcomes

We compute a set of welfare and platform-wide quantities.

- **Welfare.** Welfare on market t is the sum of consumer surplus to both unsophisticated and sophisticated agents; producer surplus; and the platform's intermediation fees.

  - **Consumer surplus to unsophisticated agents.**

$$CS_{t,U} = [(1-\rho) \times A_t]/\alpha_t \times \sum_{j \in \mathcal{J}_t} r_{jt} \ln \left[\exp(\delta_{jt}) + \exp(\delta_{0t})\right].$$

  - **Consumer surplus to sophisticated agents.**

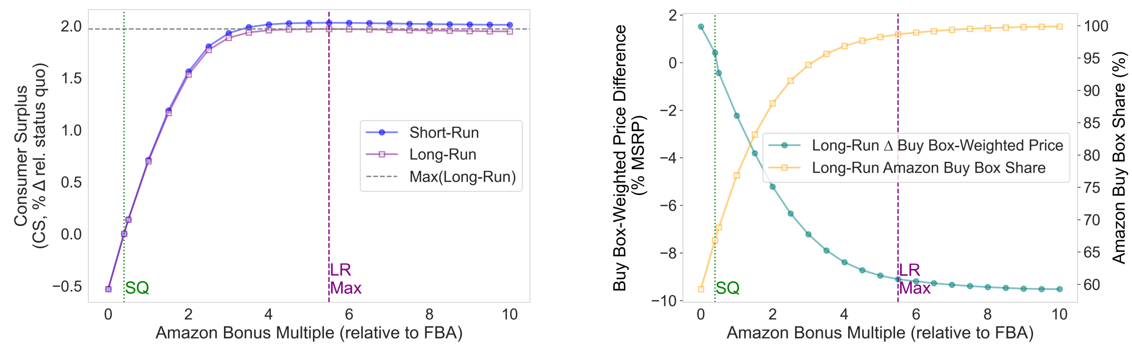
$$CS_{t,S} = (\rho \times A_t)/\alpha_t \times \ln \left[\exp(\delta_{0t}) + [\sum_{j \in \mathcal{J}_t} \exp(\delta_{jt}/\lambda)]^{\lambda}\right].$$

  - **Producer surplus.** $PS_t = A_t \sum_{j \in \mathcal{J}_t} s_{jt} \times (\phi p_{jt} - c_{jt})$.
  - **Platform intermediation fees.** $\text{Fees}_t = A_t(1-\phi) \sum_{j \in \mathcal{J}_t} s_{jt} p_{jt}$.

- **Third-party mean sales per month.** $\overline{\text{Sales}} = \frac{1}{T} \sum_{t=1}^{T} A_t \sum_{j \in \mathcal{J}_t} s_{jt}$.
- **Third-party mean price (as % of MSRP).** $\bar{p} = \frac{1}{T} \sum_{t=1}^{T} \left[\frac{1}{N_t \times \text{msrp}_t} \sum_{j \in \mathcal{J}_t} p_{jt}\right]$.
- **Third-party mean minimum price (as % of MSRP).** $\bar{p}_{\min} = \frac{1}{T} \sum_{t=1}^{T} \left[\frac{1}{msrp_t} \min_{j \in \mathcal{J}_t} p_{jt}\right]$.
- **Third-party mean cost (as % of MSRP).** $\bar{c} = \frac{1}{T} \sum_{t=1}^{T} \left[\frac{1}{N_t \times msrp_t} \sum_{j \in \mathcal{J}_t} c_{jt}\right]$.
- **Third-party mean number of entrants.** $\bar{N}_t$.

G.2 Is Amazon Self-Preferencing if Amazon Prices are Fixed?

In our paper, our leading counterfactual analyses assume that the price of the Amazon offer, where it exists, is set to maximize short-run profits. However, one could also take the view that the platform sets the price of its offers with some long-run objective in mind: for instance, as a benchmark to encourage comparable third-party offers to price more competitively or in an attempt to attract more consumers to the platform. As discussed in the main text, such a long-run objective is consistent with both criticisms of the company (Khan 2016) and statements by the company's leadership (Rose 2013).



(a) Consumer Surplus **(b)** Price & Amazon Prom. Share

Figure 10: If Amazon does not adjust its prices, how much should it guide consumers to its own offers?

Notes: The left panel shows the impact on consumer surplus as we vary the weight that Amazon's prominence algorithm places on Amazon's own offers, holding fixed the price of the Amazon offer, where it exists. The right panel shows the corresponding effects on prices and the average share of the Amazon offer in prominence, conditional on the Amazon offer being present. The vertical dashed line represents the value of $\hat{\beta}^r_{Amazon}/\hat{\beta}^r_{FBA}$ that maximizes the long-run consumer surplus, while the vertical dotted line represents the Status Quo.

Therefore, we now assess the robustness of our result that Amazon is slightly self-preferencing to the assumption that Amazon's retail division sets prices to maximize short-run profits. To this end, in Figure 10, we run the same counterfactual analysis as in Figure 5, but now we assume that the price of the Amazon offer is fixed at the Status Quo level. Under this alternative assumption, we find that consumer surplus would be maximized by a much higher prominence weight on Amazon's own offers than in Section 7.3. Indeed, we find that Amazon is insufficiently steering consumers towards its own offers in the Status Quo, and hence that the platform is not self-preferencing. This result is driven by strong consumer preferences for the Amazon offers, but also by their comparatively low prices: as we see in Figure 10b, the Buybox-weighted price falls as we increase the prominence weight on Amazon's offers.

We emphasize that the difference between exogenous and endogenous pricing – while it flips the answer to the question whether Amazon is self-preferencing in the sense of harming consumers – is smaller than it may first appear: while we do find self-preferencing in Section 7.3, we also find that the consumer surplus loss from this self-preferencing is negligible.

Appendix References

Ackerberg, Daniel A (2009). "A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation". *Quantitative Marketing and Economics* 7.4, 343–376.

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.

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

Cartis, Coralia, Jan Fiala, Benjamin Marteau, and Lindon Roberts (2019). "Improving the Flexibility and Robustness of Model-Based Derivative-Free Optimization Solvers". *ACM Transactions on Mathematical Software (TOMS)* 45.3, 1–41.

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

Conlon, Christopher and Jeff Gortmaker (2020). "Best Practices for Differentiated Products Demand Estimation with PyBLP". *RAND Journal of Economics* 51.4, 1108–1161.

Khan, Lina M (2016). "Amazon's antitrust paradox". *Yale Law Journal* 126.

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

— (1989). "A method of simulated moments for estimation of discrete response models without numerical integration". *Econometrica*, 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". *arXiv preprint arXiv:1012.5836*.

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

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

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

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.

Rose, Charlie (2013). *Amazon*. Television broadcast. 60 Minutes. Aired December 1 on CBS.