Inventory Management in Markets with Search Frictions

Fei Li Charles Murry Can Tian Yiyi Zhou*

June 8, 2021

Abstract

We present a directed search model of intermediaries with dynamic inventory and revenue management. Buyers purchase goods produced by sellers through intermediaries. Search frictions create demand uncertainty and make instantaneous replenishment impossible. To avoid the risk of stock-out, intermediaries hold inventory and employ inventory-based pricing and ordering policies. In equilibrium, when inventory is high, the intermediary posts a lower retail price to speed up sales and depresses wholesale price to slow down purchases, generating unimodal steady-state distributions of both inventory holdings and retail prices. Using a dataset that contains detailed information on used-car listings, we examine the dynamics of car dealers' inventories, new orders, sales, and retail prices and find evidence supporting the theory. We calibrate the model and show that (i) the proposed mechanism can generate a substantial amount of observed price dispersion, and (ii) our calibration attributes a substantial proportion of intermediaries' value to inventory management.

Keywords: Directed Search, Inventory Management, Revenue Management, Intermediary, Price Dispersion, Used-Car Dealers.

JEL Classification Codes: D82, D83, L15, L62

*Fei Li: Economics Department, University of North Carolina at Chapel Hill, lifei@email.unc.edu. Charles Murry: Economics Department, Boston College, charles.murry@bc.edu. Can Tian: Economics Department, University of North Carolina at Chapel Hill, tiancan@email.unc.edu. Yiyi Zhou: Department of Economics and College of Business, Stony Brook University, yiyi.zhou@stonybrook.edu.

1 Introduction

Dealers, retailers, and other intermediaries play a prominent role in well-functioning markets.[1] Although their rationale is often explained as mitigating search frictions (see, e.g. Rubinstein and Wolinsky, 1987; Gavazza, 2016), demand and supply may be uncertain because search frictions cannot be fully mitigated. Intermediaries often hold inventory as a way to manage uncertain demand, which involves (1) managing new orders and scrappage decisions, and (2) managing revenue and sales through dynamic pricing policies. We study the role of intermediaries in managing inventory, the value this behavior adds to the market, and the role that inventory management influences price dispersion.

Specifically, we introduce inventory management into an equilibrium search model, analyze equilibrium price dynamics and cross-sectional price dispersion, and calibrate the model to quantitatively evaluate the contribution of inventory-based policy to used-car dealers' welfare. The model generates price dispersion patterns consistent with real markets even though there is no ex-ante heterogeneity among buyers, sellers, or intermediaries. In particular, the model implies unimodal price distributions and patterns of inventory management consistent with recent empirical studies of price dispersion, as well as data we present on used-car dealers. We calibrate the model to used car dealer inventories and find that half of an intermediaries' value is due to holding multiple units of inventory and employing inventory-based dynamic pricing.

We borrow elements of models from the directed on-the-job search literature *á la* Menzio and Shi (2011) to model intermediaries' dynamic inventory management and pricing in the presence of market frictions in both demand and supply.[2] In our model, it takes time for buyers and sellers to meet intermediaries. Given current inventories, intermediaries decide on bid prices for future orders, the number of orders, and retail prices. Search is directed in the sense that, in retail markets, buyers observe all retail prices and decide the set of intermediaries to seek. Analogously, in wholesale markets, sellers observe all requested orders and decide the set of intermediaries to search. The model shares some flavor of Chamberlin's monopolistic competition insight: each intermediary faces a downward-sloping demand curve in retail markets and an upward-sloping supply curve in wholesale markets, but each intermediary is negligible in the sense that it can ignore its impact on, and hence reactions from, other intermediaries, making each intermediary's dynamic pricing and inventory management decision a monopolistic control problem. The tractability allows us to separately solve each agent's equilibrium policy function and the steady-state distributions of inventory and retail prices.

We solve the model for intermediaries' equilibrium retail pricing rules and stocking decisions, which depend on their current inventory size. In equilibrium, intermediaries' inventories fluctuate around a cutoff level. When inventory is above the cutoff level (e.g., because of unusually low demand), the risk of stocking out is low. The intermediary then offers low prices for new orders in

[1] According to Spulber (1996), about a quarter of U.S. economic activity is contributed by intermediaries.

[2] This tractable framework has been successfully applied in various contexts such as firm dynamics (Schaal 2017), schooling (Guo 2018), relational lending (Boualam 2019), etc. We refer readers to Wright, Kircher, Julien, and Guerrieri (2017) for a comprehensive survey of the literature.

the wholesale market and charges high retail prices. As a result, future inventory is likely to fall. When inventory is below the cutoff level, the risk of stocking out is high, so the intermediary offers high prices for new orders and a relatively low retail price, and the future inventory is likely to rise. The idiosyncratic dynamics of inventory generates *temporal price dispersion*. In steady-state, the distribution of inventory has a single-mode at the cutoff level. Because the retail price is a monotone function of the inventory, the steady-state equilibrium distribution of retail price is single-peaked. We consider multiple extensions, including a model with multiple wholesale units that lead to an optimal (s, S) rule policy for inventory acquisition.

We apply our theory to used-car dealer inventory management and pricing, which is a novel setting because data on large decentralized markets and inventory is not readily available. We have the access to data from an online used-car platform, cars.com, which includes used-car dealers' inventories and list prices. We apply our theory to the used car industry for which we have access to data that contains the weekly information of the orders, sales, inventories and list prices of all listed used cars of a large number of used-car dealers in a single U.S. state for one year. To the best of our knowledge, our paper is the first one that empirically examines intermediaries inventory and pricing in such a large decentralized market, mainly due to the unavailability of appropriate data. The used-car industry is a natural environment to study the relationship between inventory and pricing. Car dealers face inventory costs, inventory decisions do not involve long-term contracts, dealers can adjust prices quickly, the wholesale market is relatively liquid, and a majority of the tens of millions of yearly used-car sales happen through dealers.[3]

We document a number of empirical observations that find a unified explanation in our model. First, we find that a car's list price is \$19 lower on average if the dealer's inventory of similar cars is one unit higher. Moreover, the effect of inventory on retail price is stronger for large dealers than for smaller dealers. We also find that car dealers dynamically management their inventory through sales and new orders in a way consistent with our model. Specially, we estimate an acquisition inventory elasticity of -0.65 and a sale inventory elasticity of 0.85. Second, there is a substantial price dispersion among *observationally identical* used cars, and the within-dealer distribution is unimodal.

The model parameters are identified from the transition probability matrix of inventory levels and the retail prices at those inventory levels. We calibrate the model using the most popular car category in our sample – 4-6 year-old non-luxury sedans. Using the calibrated model, we quantify the role of inventory management on equilibrium prices and the level of price dispersion. The model generates a \$154 standard deviation in the retail price, and its generated 90-10 inter-percentile range is \$485. Depending on the dispersion measure and how heterogeneity is controlled, the model-generated dispersion accounts for 9%-58% of the empirical price dispersion.

[3]Although the used car market is a leading example of the classic lemons problem (Akerlof, 1970), a large literature on informational intermediaries demonstrates that, both theoretically and empirically, one of the most important functions of car dealers is to mitigate (if not fully resolve) information asymmetry between buyers and sellers. See, e.g., Biglaiser (1993), Biglaiser, Li, Murry, and Zhou (2020), and Adams, Hosken, and Newberry (2011). The basic argument is that dealers have the expertise to effectively detect lemons, and they have the incentive not to sell lemons due to the standard reputation concerns. See Biglaiser and Friedman (1994) who formalize this idea.

We use the calibrated model to investigate the value of inventory management. We find that used-car dealers' profit doubles when they are allowed to optimally manage their inventory compared to the case in which they can hold at most one unit of inventory. Furthermore, we quantify the contributions of inventory-based pricing on the retail wholesale sides of the market by shutting down the inventory-based pricing on one side. We find that retail-side inventory-based pricing is more valuable than wholesale-side management to car dealers.

Related Literature and Contribution. Our paper contributes to the literature on intermediaries that argue intermediaries mitigate search frictions. There are many theoretical studies on this topics in various settings. See, e.g., Rubinstein and Wolinsky (1987), Rust and Hall (2003), Duffie, Gârleanu, and Pedersen (2005), Wright and Wong (2014), and Nosal, Wong, and Wright (2019). This theoretical hypothesis has also been supported by some recent empirical research. See, e.g., Gavazza (2016) and Salz (2017). We recommend Gavazza and Lizzeri (2021) for a comprehensive survey of this literature. Our novelty is to introduce intermediaries' inventory and revenue management and study the implications on price dynamics and dispersion. While most papers adopt random search models, our paper explores another more tractable approach *à la* Menzio and Shi (2011) that exploits directed search and block recursivity. This tractability enables a straightforward characterization of the equilibrium inventory and price dynamics and a transparent discussion of the main trade-off of inventory management in the presence of search frictions. In addition, our model can easily accommodate rich heterogeneities and even aggregate uncertainty, making it a tractable tool for applications.

We are certainly not the first to consider inventory with search frictions. But previous studies are pure theoretical. Their focuses are the scale effect of holding multiple inventories and the corresponding benefit to the intermediaries. For instance, Johri and Leach (2002), Shevchenko (2004) and Smith (2004) introduce consumer preference heterogeneity and highlight the benefit of holding multiple units of inventory to satisfy diverse preference. In a recent paper, Rhodes, Watanabe, and Zhou (2021) introduce multiple products and study the optimal portfolio choice of intermediaries. In contrast, we investigate inventory-based dynamic pricing and ordering both theoretically and empirically. We also quantitatively evaluate the contribution of inventory-based policies to intermediaries' profit.

Also related is the literature on inventory management and pricing. While the idea to combine pricing and inventory management was first proposed by Whitin (1955), few studies have been done to understand the impact of these practices on equilibrium price dynamics and dispersion in a competitive environment. One exception is the literature that combines demand uncertainty and costly capacity *à la* Prescott (1975) to generate price dispersion and inventory holding. See, e.g., Bental and Eden (1993) and Deneckere, Marvel, and Peck (1996). The current paper differs from these models by studying the impact of search frictions in both retail and wholesale markets. To the best of our knowledge, our paper is the first one that fills the gap between the literature of dynamic inventory management and equilibrium search theory.

The negative relation between inventory and price has been documented in many markets.

Using supermarket data, Aguirregabiria (1999) highlights the role of inventory on pricing. He empirically shows that the stock-out probability and the fixed cost of ordering can explain discounts/sales behavior. Hall and Rust (2000) investigate a U.S. steel wholesaler and find that orders and sales are made infrequently and contribute to substantial variation in output prices. Zettelmeyer, Morton, and Silva-Risso (2006) document the transaction price of a new car decreases in the dealer's inventory. Also focusing on new-car markets, Copeland, Dunn, and Hall (2011) document the time series of prices, sales, and inventories and propose an inventory-based pricing model to rationalize the empirical finding.

Finally, we provide an empirically relevant new angle to rationalize price dispersion by showing that, without any ex-ante heterogeneity, our search-theoretic model of intermediates' inventory management generates a unimodal price dispersion via inventory dynamics in a pure-strategy equilibrium. It has been well-documented that observationally equivalent (after controlling observable characteristic of products and sellers) products are sold at different prices in many industries, including groceries and other traditional retail markets, wholesale markets, financial products, and illegal drugs. The theoretic literature typically (i) requires buyers with (essentially) heterogeneous information (Burdett and Judd 1983, and Stahl 1989) or sellers with heterogeneous cost or visibility (Reinganum 1979) to generate price dispersion in mixed-strategy equilibria,[4] or (ii) relies on non-stationarity of search (Coey, Larsen, and Platt 2020). See Baye, Morgan, and Scholten (2006) for a survey of the literature. Importantly, our model's prediction are robust in the sense that, regardless of the value of the parameters, the residual price distribution must be unimodal, which is consistent with many empirical studies (see, e.g., Kaplan and Menzio 2015).

Organization. The rest of the paper is organized as follows. Section 2 introduces the theoretical model. Section 3 characterizes the equilibrium and derives empirical implications. We discuss some simple extensions of the base model in section 4. Section 5 describes the dataset, constructs the working sample, and documents how the intermediaries' new orders, sales, and retail price relate to their inventory level. It also provides evidence that the inventory and retail price are unimodally distributed after controlling for other confounding heterogeneity. Section 6 calibrates the model using the data of the used-car markets. Section 7 concludes. Omitted proofs are relegated to Appendix A, additional empirical results are presented in Appendix B, and additional quantitative results are in Appendix C.

2 Model

2.1 Environment

We consider a continuous-time model with infinite horizon. The economy is populated by buyers, sellers and intermediaries.

[4]The qualification "essentially" is added because the heterogeneous information structure can be endogenized by adding a stage of costly information acquisition of homogeneous consumers as in Burdett and Judd (1983).

Buyers and Sellers. At each instant, a continuum of potential buyers and sellers is present. They are short-lived. Each seller has a unit supply of the indivisible consumption good, and he receives zero utility by consuming the good by himself. Each buyer has a unit demand of the consumption good, and by consuming the good, his utility is $u > 0$. Despite the gain from trade, buyers and sellers face some obstacles to trade, creating a role of intermediaries. Our focus is to model the equilibrium mediated transaction mechanism where the consumption goods are delivered from sellers to buyers through intermediaries.

Intermediaries. There is a unit measure of ex-ante identical long-lived intermediaries (dealers), each of whom purchases consumption goods from sellers in the wholesale market and sells consumption goods to buyers in the retail market. An intermediary can sell only if his current inventory is positive. To avoid the risk of stockout, an intermediary can hold inventory. The flow cost of holding x units of inventory is $c(x)$ for $x = 0, 1, 2, \ldots$. The cost function $c : \mathbb{N} \to \mathbb{R}$ is increasing, weakly convex, and such that $c(0) = 0$. All intermediaries are risk-neutral and share a common discount rate $\rho > 0$.

Let $g_t : \mathbb{N} \to [0, 1]$ be the probability mass function of the distribution of inventory holding across intermediaries at time t. Specifically, $g_t(x)$ represents the measure of intermediaries who holds x units of inventory at time t. Therefore, $g_t(x) \geq 0, \forall t, x$ and $\sum_{x \in \mathbb{N}} g_t(x) = 1, \forall t$.

Markets. The retail market is organized in multiple submarkets indexed by the retail price $p \in \mathbb{R}$. In each retail submarket p, the ratio of buyers to intermediaries is denoted by $\theta(p)$. Retail submarket p can therefore be viewed as a group of agents who wish to trade at price p. Similarly, the wholesale market is organized in multiple submarkets indexed by the wholesale price $w \in \mathbb{R}$. In each wholesale submarket w, the ratio of sellers to intermediaries is denoted by $\lambda(w)$. Following Pissarides (1985), we refer to $\theta(p)$ and $\lambda(w)$ as the *tightness* of the corresponding retail and wholesale submarkets.

Search and Matching. Search is directed. At each moment, an intermediary can choose to enter exactly one retail submarket and one wholesale submarket simultaneously. In this way, we capture the intermediary's retail/wholesale pricing problem as a choice of the corresponding submarkets. A buyer sees all the retail submarkets (prices) and chooses to enter at most one retail submarket to search for intermediaries at each time. The flow search cost to stay in a retail submarket is $\kappa_b > 0$.[5] Similarly, a seller sees all wholesale submarkets (prices) and chooses to enter at most one wholesale submarket at each moment. It costs the seller $\kappa_s > 0$ at each moment by staying in a wholesale submarket. When an intermediary and a buyer meet in retail submarket p, the buyer buys one unit of the good from the intermediary at price p. When an intermediary and a seller meet in wholesale submarket w, the seller sells one unit of the good to the intermediary at price w.

[5]The assumption that a buyer can visit at most one submarket at each time can be relaxed. Alternatively, one can allow a buyer to visit n submarkets by incurring a cost $n\kappa_b$ where n is some positive integer and can be either exogenously specified or endogenously chosen *à la* Stigler (1961). In this case, the rate at which the buyer meets an intermediary is proportional to n as well.

There are frictions in submarkets. In each submarket, the matching process is determined by a matching function. Following Pissarides (1985), we assume that the matching function is homogeneous of degree one so that the matching process in each submarket is fully determined by its tightness. Specifically, at each instant, an intermediary meets a buyer at a rate $\phi_r(\theta(p))$ in retail submarket p. We further assume that $\phi_r : \mathbb{R}_+ \to \mathbb{R}_+$ is a bounded, twice-differentiable, strictly increasing and strictly concave function such that $\phi_r(0) = 0$. On the other side, a buyer makes a contact with an intermediary at a rate $\psi_r(\theta(p))$ where $\psi_r : \mathbb{R}_+ \to \mathbb{R}_+$ is also a bounded, twice-differentiable, strictly decreasing and strictly concave function such that $\lim_{\theta\to\infty} \psi_r(\theta) = 0$. As in Pissarides (1985), we have $\psi_r(\theta) = \phi_r(\theta)/\theta, \forall\theta > 0$. This is because the number of intermediaries who meet buyers must equal the number of buyers who meet intermediaries.

Similarly, in a wholesale submarket w, an intermediary meets a seller at rate $\phi_w(\lambda(w))$ where $\phi_w : \mathbb{R}_+ \to \mathbb{R}_+$ is a bounded, twice-differentiable, strictly increasing and strictly concave function such that $\phi_w(0) = 0$. On the other side, a seller meets an intermediary at rate $\psi_w(\lambda(w))$ where $\psi_w : \mathbb{R}_+ \to \mathbb{R}_+$ is a bounded, twice-differentiable, strictly decreasing function such that $\lim_{\lambda\to\infty} \psi_r(\lambda) = 0$, and $\psi_w(\lambda) = \phi_w(\lambda)/\lambda, \forall\lambda > 0$.

Discussion of Assumptions. Before moving forward, we discuss some assumptions. First, we assume that search is directed. An agent is fully aware of the price and the matching probability of each submarket. This specification is not essential. What matters is that it both captures search friction and also preserves the familiar trade-off between transaction speed and price in a competitive environment. Specifically, to sell faster, the intermediary has to enter a retail submarket with higher tightness, implying a lower equilibrium retail price. Similarly, to buy faster, the intermediary has to enter a wholesale submarket with higher tightness, implying a higher equilibrium wholesale price. One can obtain a similar trade-off in a random search model by introducing random utility (demand curve) and production cost (supply curve).

Second, we assume homogeneous product and ex-ante homogeneous intermediary. The first assumption is made to emphasize the search frictions resulting from the uncertainty about how quickly agents are matched. This is a deliberate simplification to highlight our main mechanism. In the literature of labor and consumer search, another important source of search frictions comes from the uncertainty regarding the match quality between agents and products, which can also be accommodated by a straightforward extension of our model (see section 4.2). We assume that intermediaries are ex-ante homogeneous to highlight the contribution of the ex-post inventory dynamics on price dispersion. This leads to a common optimal inventory-based pricing policy among intermediaries. It is straightforward to extend our model to allow for ex-ante heterogeneous inventory-price relationship (see section 4.3).

Third, we assume that buyers and sellers are short-lived. This assumption implies that there is no long-term relationship between buyers and intermediaries or sellers and intermediaries. Also, as we shall show, the equilibrium measure of active buyers and sellers is pinned down by free-entry conditions. These assumptions allow us to focus on the dynamic problem of intermediaries and to maintain tractability. In our partial-equilibrium model, one can think of the entry cost of sellers and search cost of buyers as their outside option by searching for each other in another frictional market. It is not difficult to extend our model to accommodate unmediated transactions and consider a general-equilibrium model where at each moment, a fixed measure of short-lived buyers and sellers decide whether to enter an unmediated frictional market to search for each other or enter a mediated submarket to search for intermediaries. One can view κ_b and κ_s of our baseline model as a reduced-form parametrization of the equilibrium payoffs of buyers and sellers in this general-equilibrium model. .

2.2 Individual Problems and Definition of Equilibrium

We solve for equilibria where agents' decisions are independent of the distribution over the inventory holding g_t. We then argue that such a focus is without loss of any generality.

The Intermediary's Problem. Consider an intermediary with inventory x. His lifetime expected profit $V(x)$ obeys the following Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{aligned}\rho V(x) = \max_{p,w\geq 0} -c(x) \quad &+ \quad \underbrace{\phi_r(\theta(p))[p + V(x-1) - V(x)]}_{\text{retail problem}} \\ &+ \quad \underbrace{\phi_w(\lambda(w))[-w + V(x+1) - V(x)]}_{\text{wholesale problem}},\end{aligned} \tag{1}$$

for every $x > 0$. At each moment, the intermediary chooses a retail submarket and a wholesale submarket to enter. There are three terms on the right-hand side of (1). The first term is the flow cost of holding x units of inventories. The second term is the rate at which the intermediary meets a buyer, $\phi_r(\theta(p))$ in retail submarket p, times the change in the intermediary's continuation value when the buyer purchases one unit of good from the intermediary, $p + V(x-1) - V(x)$. The third term is the rate at which the intermediary meets a seller, $\phi_w(\lambda(w))$ in wholesale submarket w, times the change in his continuation value when he buys one unit of good from the seller. In words, an intermediary takes the tightness of each submarket as given and controls the (Poisson) arrival rates of buyers ($\phi_r(\theta(p))$) and sellers ($\phi_w(\lambda(w))$) by choosing retail and wholesale prices at each movement.[6]

An intermediary cannot sell when stockout ($x = 0$), so

$$\rho V(0) = \max_{w\geq 0} \phi_w(\lambda(w))[-w + V(1) - V(0)].$$

The Problem of Buyers and Sellers. At each instant, a buyer decides whether and where to

[6]As it is standard in the literature, the continuous-time Poisson arrival rate setting relieve us from specifying the event where the intermediary finds a buyer and a seller at the same time. Also note that while the intermediary's retail and wholesale decisions are expressed in separated terms in equation (1), they must be made jointly as the value function is determined by both retail and wholesale policies.

search. So the tightness of each retail submarket must satisfy

$$\kappa_b \geq \psi_r(\theta(p))(u-p), \tag{2}$$

and $\theta(p) \geq 0$ with complementary slackness. Condition (2) guarantees that the tightness $\theta(p)$ is consistent with the consumer's incentive to search. The cost of search is given by κ_b, and the benefit of search is given by the product between the rate at which the consumer meets an intermediary $\psi_r(\theta(p))$ and the surplus from buying the good at price p. Moreover, if the tightness is strictly positive, we say the submarket is *active*, and the complementary slackness condition implies that the search cost must equal the benefit. It is obvious that in any active retail submarket, $p < u$. On the other hand, if the submarket is inactive ($\theta(p) = 0$), the search cost can be either greater than or equal to the benefit. Similarly, the tightness of each wholesale submarket must satisfy

$$\kappa_s \geq \psi_w(\lambda(w))w, \tag{3}$$

and $\lambda(w) \geq 0$ with complementary slackness. The left-hand side of condition (3) is the entry cost κ_s; while the right-hand side corresponds the expected revenue of entry, which is given by the product between the rate at which the seller meets an intermediary $\psi_w(\lambda(w))$ and the selling price of the good w. Whenever $\lambda(w) > 0$, the entry cost equals the expected revenue that ensures the standard zero-profit condition, and $w > 0$. On the other hand, if $\lambda(w) = 0$, the entry cost is greater than or equal to the revenue.

Definition of Equilibrium. A *block recursive equilibrium* consists of two market tightness functions, $\theta : \mathbb{R} \to \mathbb{R}$ and $\lambda : \mathbb{R} \to \mathbb{R}$, a value function for the intermediary, $V : \mathbb{N} \to \mathbb{R}$, a retail pricing policy $p : \mathbb{N} \to \mathbb{R}$, and a wholesale pricing policy $w : \mathbb{N} \to \mathbb{R}$. These functions satisfy the following conditions:

1. $V(x)$ and $p(x), w(x)$ solve problem (1), and
2. the retail market tightness function $\theta(p)$ satisfies condition (2), and wholesale market tightness function $\lambda(w)$ satisfies condition (3).

Our solution concept, developed by Shi (2009) and Menzio and Shi (2010, 2011), is similar to the standard recursive competitive equilibrium except that the agent's problem does not depend on the distribution of heterogeneous individual states across agents g. In general, the aggregate state g varies over time, and it is natural to conjecture that the dynamics of g affects intermediaries' value functions and optimal decisions. However, thanks to the insight of Menzio and Shi (2010, 2011), in our setting, it is without loss of generality to focus on block recursive equilibria because all recursive equilibria are block recursive.[7] While the standard formal argument of the uniqueness is omitted, the intuition will be further elaborated later.

[7]It is worthwhile noting that the distribution-free property of our *competitive* equilibrium solution concept also relieves us from the usual belief-based multiplicity in most game-theoretic two-sided market models, for example in Caillaud and Jullien (2003) and the subsequent literature.

3 Analysis

3.1 Equilibrium Characterization

Because search is directed, an intermediary faces a trade-off between the trade probability and the transaction price. Specifically, the optimality condition of consumer search (2) implies that the equilibrium expected benefit of search must be constant for every active retail submarket. Hence, $\theta(p)$ must decrease in p. That is, if a retail submarket features a higher price, its equilibrium buyer-to-intermediary ratio must be lower, making it more likely for each consumer to meet an intermediary. On the other side of the coin, if an intermediary wants to sell at a higher speed (larger $\phi_r(\theta(p))$), he must enter a retail submarket featuring a lower price p. The same reasoning applies to the trade-off between order speed and price in wholesale submarkets. In the next paragraphs, we formalize the arguments.

By condition (2), it must hold that in every active retail submarket, one can represent the retail price p as a function of its tightness $\theta(p)$:

$$p(\theta) = u - \frac{\kappa_b}{\psi_r(\theta)} = u - \frac{\kappa_b \theta}{\phi_r(\theta)}, \tag{4}$$

where the second equality holds because $\phi_r(\theta) = \psi_r(\theta)/\theta, \forall \theta > 0$. Because $\psi_r(\cdot)$ is decreasing, condition (4) immediately implies that $p(\cdot)$ is decreasing. It can be viewed as the "demand curve" an intermediary faces. Similarly, condition (3) implies that, in every active wholesale submarket, the wholesale price w and the tightness $\lambda(w)$ are such that

$$w(\lambda) = \frac{\kappa_s}{\psi_w(\lambda)} = \frac{\kappa_s \lambda}{\phi_w(\lambda)}, \tag{5}$$

and w is increasing in λ, which has the flavor of the "supply curve" faced by an intermediary.

Substituting conditions (4) and (5) into the intermediary's Bellman equation (1), we can express the intermediary's pricing problem as

$$\begin{aligned} \rho V(x) = -c(x) \quad &+ \quad \max_{\theta \geq 0} \underbrace{\phi_r(\theta)[u + V(x-1) - V(x)] - \kappa_b \theta}_{\text{retail transaction gain}} \\ &+ \quad \max_{\lambda \geq 0} \underbrace{\phi_w(\lambda)[V(x+1) - V(x)] - \kappa_s \lambda}_{\text{wholesale transaction gain}}. \end{aligned} \tag{6}$$

That is, the intermediary's pricing problem is equivalent to a problem where the intermediary chooses the tightness of the retail submarket where he looks for buyers and the tightness of the wholesale submarket where he looks for sellers. The optimal retail market tightness, denoted by $\theta^*(x)$, maximizes the expected surplus generated by a transaction between the intermediary and a buyer, net of the consumer search cost per intermediary to maintain the market tightness to be $\theta^*(x)$. The value of the optimization problem corresponds to the retail transaction gain to the intermediary. Similarly, the optimal wholesale market tightness, denoted by $\lambda^*(x)$, maximizes the expected surplus generated by a transaction between the intermediary and a seller, net of the seller's entry cost to maintain the market tightness to be $\lambda^*(x)$. The corresponding value of the optimization problem captures the wholesale transaction gain to the intermediary. Therefore, solving the equilibrium is equivalent to solving decision problem (6).

The above argument also makes it clear why the distribution over inventory holding across intermediaries is payoff irrelevant in equilibrium. An intermediary wants to maximize the expected life-time total utility that he delivers to buyers, net of the expected life-time inventory cost. The probability distribution over the path of his future utility creation and inventory cost are solely pinned down by the endogenous choices θ and λ in subsequent periods. Therefore, the distribution g has nothing to do with the intermediary's continuation payoff. By conditions (4) and (5), the equilibrium retail and wholesale prices in each submarket are solely pinned down by the corresponding tightness θ and λ, and so they are also independent of the aggregate state g. The reasoning crucially depends on the assumption that search is directed, allowing intermediaries with different inventory sizes to trade at different prices at different speeds (in different submarkets). If search is random, corresponding to the setting with one retail submarket and one wholesale submarket, then the trade-off between the price and the transaction speed disappears without further twisting of the current setting. The discussion above is summarized by the following proposition. The formal argument is essentially identical to the one in Menzio and Shi (2010, 2011) and Li and Weng (2017), so it is omitted.

Proposition 1. *A (block recursive) equilibrium exists and it is unique.*

Now we are ready to characterize the unique equilibrium. From problem (6), it follows that the optimal $\theta^*(x)$ must satisfy the following first-order condition (FOC):

$$\kappa_b \geq \phi_r'(\theta^*(x))[u + V(x-1) - V(x)], \tag{7}$$

and $\theta^*(x) \geq 0$ with complementary slackness. The FOC in condition (7) says that the social marginal cost to maintain the tightness to be $\theta^*(x)$ must equal the social marginal benefit of doing so in any active retail submarket. Here, the social cost is incurred by buyers and the social benefit is generated by the expected surplus of a transaction. Similarly, the optimal $\lambda^*(x)$ must satisfy

$$\kappa_s \geq \phi_w'(\lambda^*(x))[V(x+1) - V(x)], \tag{8}$$

and $\lambda^*(x) \geq 0$ with complementary slackness. It says that the social marginal cost incurred by sellers equals the social marginal benefit of maintaining the tightness to be $\lambda^*(x)$. The discussion above implies that the equilibrium allocation is socially efficient.

Notice that conditions (7) and (8) imply that $\theta(x)$ and $\lambda(x)$ depend on the gain from trade, $u + V(x-1) - V(x)$ and $V(x+1) - V(x)$, respectively, which depends on the intermediary's current inventory size x. The following lemma characterizes how inventory size affects the gain from trade for an intermediary in both retail and wholesale markets.

Lemma 1. *In the equilibrium, $V(x+1) - V(x)$ decreases in x.*

Lemma 1 says that the marginal benefit of accumulating inventory is decreasing. We provide the intuition here and relegate the proof to Appendix A. The benefit of holding inventory is to lower the risk of stockout. In the presence of search frictions, an intermediary faces uncertainty about both the demand in retail markets and the supply in wholesale markets. If his inventory is reduced to zero, he can neither immediately order goods from sellers nor trade with buyers. The risk of stockout strictly decreases in the intermediary's inventory size x; so does the marginal benefit of holding inventory, making the dealer's value function concave. Notice that the concavity of the intermediary's value function reflects the diminishing of stockout risk. The convexity of inventory cost function certainly contributes to the concavity of the value function, but it is not necessary. The value function is concave even in the absence of inventory cost or when the inventory cost function is affine.

Now we are ready to derive the relationship between inventory and prices. Let

$$p^*(x) \equiv p(\theta^*(x)) \text{ and } w^*(x) \equiv w(\lambda^*(x))$$

denote the equilibrium retail and wholesale pricing policy where $p(\cdot)$ and $w(\cdot)$ are specified in conditions (4) and (5).

Proposition 2. *In the equilibrium, the intermediary's choice of submarkets is such that*

1. *$\theta^*(x)$ increases in x and retail price $p^*(x)$ decreases in x and*
2. *$\lambda^*(x)$ and the wholesale price $w^*(x)$ decrease in x.*

Proposition 2 says that when his inventory increases, the intermediary will enter a retail submarket with lower price and higher matching probability (easier to sell), and enter a wholesale submarket with lower wholesale price and lower matching probability (harder to buy). This is intuitive. An intermediary trades off between the risk of stockout and the cost of inventory and new orders. When the inventory stock becomes higher, the risk of stockout decreases, but the inventory cost becomes higher, so it is optimal to lower future inventory by selling more and buying less. To do so, the intermediary needs to lower both the retail and wholesale prices. Similarly, when his stock becomes too low, the concern of stockout grows, and the intermediary raises both the retail price and the wholesale price to slow down the sales and speed up new orders, increasing his future inventory holding in expectation.

The empirical implication of Proposition 2 is that when the intermediary's inventory increases, (i) the retail price decreases and the sales increase on average, and (ii) the wholesale price decreases and new orders decrease on average.

Corollary 1. *An intermediary's retail and wholesale prices co-move over time.*

Corollary 1 is an immediate implication of Proposition 2. Driven by the change in inventory, an intermediary's retail price and wholesale price should move in the same direction. Depending on the elasticity of the matching functions in retail and wholesale markets and the search and entry cost, the retail price and the wholesale price may respond to the inventory change differently.

When the wholesale price is more sensitive to the change of inventory, the equilibrium exhibits *incomplete pass-through* (Nakamura and Zerom 2010). Also, because of the co-movement, the *markup*, which is the difference between the retail price and the wholesale price, can be either positively or negatively correlated with the inventory.

Furthermore, suppose that the inventory cost is unbounded, i.e., $\lim_{x\to\infty} c(x) = \infty$. Then $V(x)$ must be decreasing for sufficiently large x. Denote by $S \in \mathbb{N}$ such that $V(x)$ is increasing if and only if $x \leq S$. Moreover, even if $x \leq S$, the marginal benefit of increasing inventory may be sufficiently small so that

$$\kappa_s > \phi_w'(\lambda)\delta[V(x+1) - V(x)],$$

for any λ, making it impossible to generate gain from trade in the wholesale market. In this case, it is still optimal to set $\lambda = 0$. We denote by

$$s = \max\{x \in \mathbb{N} : \lambda^*(x) > 0\}, \tag{9}$$

in the equilibrium, which is referred as the *base level of stock* in the literature. Notice that $s \leq S$, and $\lambda^*(x) > 0$ for any $x \leq s$. Therefore, the equilibrium resembles the classic *based stock policy* in the inventory management literature (see, e.g., Porteus 2002).

Corollary 2. *In the equilibrium, the intermediary employs a based stock policy, i.e., $\lambda^*(x) > 0$ if and only if $x \leq s$.*

One can add the option of free disposal to keep the intermediary's value function monotone. Then, no intermediary holds inventory above S, and no intermediary orders inventory if $x > s$.

3.2 Steady-State Distribution

Now we study the steady-state distribution of inventory holding and retail price. In the equilibrium, the distribution of inventory across intermediaries evolves as follows:

$$\begin{aligned} \dot{g}_t(x) &= \underbrace{g_t(x-1)\phi_w(\lambda^*(x-1)) + g_t(x+1)\phi_r(\theta^*(x+1))}_{\text{inflows}} \\ &- \underbrace{g_t(x)[\phi_r(\theta^*(x)) + \phi_w(\lambda^*(x))]}_{\text{outflows}}, \end{aligned} \tag{10}$$

for every $x = 0, ..., s$. The left-hand side of (10) is the time derivative of the measure of intermediaries who hold x units of inventory at time t. The right-hand side of (10) has three parts. The first two terms are positive, but the last term is negative. First, $g_t(x-1)$ of intermediaries hold $x-1$ units of inventory and search in a wholesale submarket with tightness $\lambda^*(x-1)$ at time t, and $\phi_w(\lambda^*(x-1))$ of them find sellers, trade, and increase their stock to x. Second, $g_t(x+1)$ of intermediaries hold $x+1$ units of inventory and search in a retail submarket with tightness $\theta^*(x+1)$ at time t, and $\phi_r(\theta^*(x+1))$ of them find buyers, trade, and decrease their stock to x. Finally, $g_t(x)$ of intermediaries hold x units of inventory at time t and $\phi_r(\theta^*(x))$ of them meet buyers and

$+\phi_w(\lambda^*(x))$ of them meet sellers, changing their inventory from x to $x-1$ and $x+1$ respectively.

At steady state, $\dot{g}_t(x) = 0$ for every x, and so the distribution of inventory holding across intermediaries is constant over time. Denote it by g_{ss}. It must satisfy

$$[\phi_r(\theta^*(x)) + \phi_w(\lambda^*(x))]g_{ss}(x) = \phi_w(\lambda^*(x-1))g_{ss}(x-1) + \phi_r(\theta^*(x+1))g_{ss}(x+1). \tag{11}$$

Proposition 3. *There exists a unique steady-state distribution of inventory holdings across intermediaries, and it is unimodal.*

Proposition 3 says that the probability mass function g_{ss} has a single peak.[8] The intuition behind is very simple. By Proposition 2, in equilibrium, an intermediary's expected increment of inventory is decreasing in his current inventory. Therefore, there exists a *cutoff inventory level* denoted by x^* such that the intermediary's expected increment is negative if and only if his current inventory stock is above x^*. As a result, the equilibrium inventory dynamics behaves as if a "mean" regression process: Whenever an intermediary's inventory deviates from x^*, he adjusts the retail or wholesale policy θ and λ to push the future stock back to x^*. The more the stock deviates from the mean level, the faster the speed of the regression is. In the state steady, the mass at the cutoff level of inventory x^* is the highest, and the mass monotonically decreases as the stock becomes farther and farther away from x^*. As a consequence, x^* is the unique mode of the steady-state distribution.

Because the intermediary retail price is monotone in his inventory size (Proposition 2), it is immediate the equilibrium inventory dynamics shapes the steady-state distribution of retail price.

Corollary 3. *There exists a unique steady-state distribution of retail prices across intermediaries, and it is unimodal.*

That is, our model predicts that the distribution of retail price in the steady state is single-peaked. Because the inventory is most likely to be around x^*, one should expect that the intermediary's retail price is equal to or close to $p^*(x^*)$ most of the time. Extremely high or low prices will be observed rarely. As we discussed in the introduction, this is consistent with most empirical studies about price dispersion in a variety of markets where intermediaries are present. Notice that our model has no ex-ante heterogeneity among buyers, among sellers, or among intermediaries. The retail price dispersion is generated even if no agent randomizes, which distinguishes our model from most search models that rely on agents' heterogeneity and mixed-strategy to generate price dispersion.

We want to point out that at the steady state, an individual intermediary's price still changes over time due to inventory changes. Therefore, the equilibrium price exhibits *intra-distribution dynamics*. That is, the rank of an intermediary's price varies over time within the price distribution. This is because we assume that intermediaries are identical, so the model only generates

[8] Following Hartigan and Hartigan (1985), we say a probability distribution is unimodal (or single peaked) if there is a mode x^* such that the cumulative density function of the probability distribution is convex for $x \leq x^*$ and concave for $x \geq x^*$.

a temporal price dispersion rather than a "spatial" or persistent price dispersion across intermediaries. This is consistent with a number of empirical studies such as Lach (2002) and Chandra and Tappata (2011). In the literature, such a phenomenon is often used to support the mixed-strategy pricing equilibrium suggested by consumer search models. Our result suggests that, to test whether firms play mixed strategies (at least in industries where inventory costs and stockout risks are non-trivial), one may also need to take into account their inventory dynamics.

4 Extensions and Discussion

In this section, we enrich our stylized model by introducing intermediary heterogeneity, multi-unit wholesale package, and product differentiation. These extensions make our model applicable to many markets and demonstrate the model can address some questions that are usually studied in static frameworks. We demonstrate that the inventory-price relationship is robust to these perturbation, but these complications may affect the shape of retail price distribution.

4.1 Multi-Unit Wholesales and the Optimality of (s, S)-Rule

In many industries, it is reasonable to assume that an intermediary can purchase multiple units when he meets a seller. Our framework can easily incorporate this feature. Suppose that a wholesale submarket is indexed by a bundle $(w, y) \in \mathbb{R} \times \mathbb{N}$ where y is the number of product of the bundle and w is the price. For simplification, we still assume that the marginal production cost of the seller is zero, so the free-entry condition (3) remains unchanged. An intermediary therefore decides not only the wholesale purchase price but also the wholesale purchase quantitive y by choosing a wholesale submarket, so his problem becomes

$$
\begin{aligned}
\rho V(x) = -c(x) \quad &+ \quad \max_{\theta \geq 0} \phi_r(\theta)[u + V(x-1) - V(x)] - \kappa_b \theta \\
&+ \quad \max_{\lambda \geq 0, y \in \mathbb{N}} \phi_w(\lambda)[V(x+y) - V(x)] - \kappa_s \lambda.
\end{aligned} \tag{12}
$$

The optimal $\theta^*(x)$ still satisfies condition (7), but the optimal $y^*(x), \lambda^*(x)$ satisfy

$$
\kappa_s \geq \psi_w(\lambda^*(w))[V(x + y^*(x)) - V(x)].
$$

The rest of the equilibrium analysis is straightforward. What's interesting is that the model can generate the classic (s, S)-rule (Scarf 1960). Under this policy, the intermediary brings the level of inventory after ordering up to some level S if the initial inventory level x is below some level s where $s \leq S$. However, in Scarf (1960), the cost to make an y-unit order is assumed to be $w_0 + yw_1$ where w_0 is the fixed cost of making an order and w_1 is the linear unit price. In our model, we do not impose any structure on the wholesale pricing rule. Instead, a highly non-linear price rule is allowed in the equilibrium analysis. The idea is as follows. The intermediary's value function is concave. When $x \geq S$, $V(\cdot)$ is decreasing so it is not optimal to add inventory. However, when

$x = S-1$, the gain from trade $V(S) - V(S-1)$ is still very small. When the entry cost κ_s is sufficiently high, it is suboptimal to set $\lambda > 0$. The gain from trade $V(S) - V(x)$ is decreasing in the inventory level x. When x is sufficiently small, it is optimal to set $\lambda^*(x) > 0$. Also, because the marginal product cost is zero, it is immediate that $y^*(x) = S - x$.

The option of making a multiple-unit order decreases the ergodic probability of inventory being sufficiently low, shifting the peak of retail price distribution downward. It implies that there is a low average price that being charged most frequently by intermediaries. The retail price gouging occasionally emerges when there is a shortage of inventory due to unexpected high peaks or the failure of wholesale ordering for a long period.[9]

4.2 Idiosyncratic Utility

Buyers have heterogeneous tastes. A standard practice to capture this heterogeneity is to introduce (horizontally) idiosyncratic utility into the model as in Wolinsky (1986) and Anderson and Renault (1999): a buyer's payoff by consuming a product is a random variable $\tilde{u}$ with a CDF G. Suppose the buyer-product match-specific utility is independently and identically distributed across buyers and products. When a buyer and an intermediary meet, the buyer will pick his favorite product that delivers positive payoff. Therefore the buyer's free-entry condition (2) becomes

$$\kappa_b \geq \psi_r(\theta(p)) \int \max\{\tilde{u} - p, 0\} dG(\tilde{u})^x. \tag{13}$$

For simplicity, assume that the match between the buyer and the product is good and the buyer receives utility u by consuming the product with probability α, and the match is bad, and the utility is 0 with complementary probability. When a buyer meets an intermediary with inventory x, the buyer finds a good match with probability

$$\Phi(x) = 1 - (1-\alpha)^x,$$

which is strictly increasing and concave in x. Therefore, holding a large number of inventory endows the intermediary another advantage: reducing the possibility of *mismatch*. In this case, a retail submarket is indexed by (p, x), the price and the inventory size of the intermediaries who trade in this market. In equilibrium, a match between a buyer and an intermediary will lead to a transaction if only if the match is good, or the gain from trade is strictly positive. Therefore, the intermediary's problem (6) becomes

$$\begin{aligned} \rho V(x) = -c(x) \quad &+ \quad \max_{\theta \geq 0} \phi_r(\theta)\Phi(x)[u + V(x-1) - V(x)] - \kappa_b \theta \\ &+ \quad \max_{\lambda \geq 0} \phi_w(\lambda)[V(x+1) - V(x)] - \kappa_s \lambda. \end{aligned} \tag{14}$$

[9]Such a situation may appear due to natural disaster such as hurricanes. For example, see `https://www.businessinsider.com/price-gouging-in-texas-gas-prices-hurricane-2017-9`.

The optimal $\theta^*(x)$ must satisfy

$$\kappa_b \geq \phi_r'(\theta^*(x))\Phi(x)[u + V(x-1) - V(x)], \tag{15}$$

and the optimal $\lambda^*(x)$ still satisfies (8). Because $\Phi(x)$ is strictly increasing and concave in x, one can verify that $V(x)$ is still concave, and the optimal $\theta^*(x)$ is still increasing in x. This is because when x is higher, each match between a buyer and the intermediary will more likely lead to a transaction, so it is socially optimal to let more buyers search.

On the buyer's side, the tightness of each retail submarket must satisfy

$$\kappa_b \geq \psi_r(\theta(p))\Phi(x)(u - p), \tag{16}$$

and $\theta(p) \geq 0$ with complementary slackness, so the equilibrium price in each active retail submarket is given by

$$u - \frac{\kappa_b}{\psi_r(\theta^*(x))\Phi(x)}.$$

In the equilibrium, $\psi_r(\theta^*(x))$ is decreasing in x while $\Phi(x)$ is increasing in x, so the retail price may no longer be monotone in the inventory x. The intuition is as follows. When his inventory increases, the intermediary wants to sell faster, so he enters a retail submarket with higher θ. From the perspective of buyers, it is less likely to meet an intermediary in a submarket with higher θ, but conditional on meeting an intermediary, it is more likely to find a desired product due to the intermediary's larger inventory size. Therefore, the effective matching probability $\psi(\theta)\Phi(x)$ and therefore the buyer's willingness to pay may not be monotone in x in the equilibrium. Although the steady-state distribution of inventory g_{ss} is still unimodal, the distribution of retail price may not be. In our numerical examples, we do find non-monotone relationship between the equilibrium retail price and inventory and the retail price dispersion has multiple modes.

On the seller's side, the equilibrium price in each active wholesale submarket is still given by (5). The rest of the equilibrium analysis will be similar to the one of the baseline model.

4.3 Heterogeneous Intermediaries

Intermediaries may be heterogeneous in their inventory costs and matching technologies. For example, some intermediaries have outstanding marketing and sales managers, bringing them higher visibility to buyers; some intermediaries have effective purchasing departments and maintain good relationship with manufacturers, allowing them to be part of a more efficient supply chain; some intermediaries have outstanding transportation or handling and teams or low opportunity cost of the money, admitting lower marginal inventory cost. These heterogeneities can lead to different average inventory size, sales and profitability, and therefore, different inventory-pricing relationship among intermediaries.

It is straightforward to extend our model by allowing heterogeneous intermediaries. Specifically, there are J types of intermediaries, and the proportion of each type j is denoted by f_j.

Denote $c_j(\cdot)$, $\phi_{r,j}(\cdot)$ and $\phi_{w,j}(\cdot)$ as type-j intermediaries' inventory cost, retail matching probability and wholesale matching probability. A retail submarket is indexed by p, j; while a wholesale submarket is indexed by w, j. Accordingly, the market tightnesses are $\theta(p, j)$ and $\lambda(w, j)$. Each type-j intermediary solves the type-specific dynamic optimization problem as in problem (6) but with its own matching probabilities and inventory cost. Within each type, the steady-state distribution of inventory and retail price are single-peaked. The shape of aggregate steady-state distribution of inventory and retail price will depend on the distribution over types.

5 Used-Car Markets and Empirical Evidence

In the rest of the paper, we apply the model to study used-car dealers' inventory management and dynamic pricing by using detailed information on used-car listings (inventories) by a large number of car dealers. While the practice of inventory management plays out in many real-world settings, several factors make the used-car market suitable for our study. First, as Gavazza, Lizzeri, and Roketskiy (2014) argue, the market is highly decentralized, making the search and matching frictions non-trivial. As a result, many transactions are intermediated. Nationally, about two-thirds of used-car sales are made by dealers. Second, inventory management is important for used-car dealers. In general, dealers must manage both value erosion as assets age and holding costs, which include floor-plan inventory investment and cost of capital.[10] Third, cars are durable goods. Most buyers and sellers do not make frequent purchases, so it is uncommon to sign long-term contracts with dealers. Fourth, stocking decisions can be made frequently. Dealers typically acquire used cars from individuals or at wholesale auctions. Dealers may have access to multiple auctions a week at multiple auction locations. Fifth, dealers frequently adjust prices. These features suggest that the interaction between inventory control and search friction is important in the used-car market, making our theory applicable.[11]

The goal of this section is twofold. First, we report empirical evidence supporting the predictions regarding inventory-based pricing and order from the used-car market. Recall that the theoretical model resembles two robust predictions from the theoretic literature of inventory management: (i) sales increase in inventory and new orders decrease in inventory, and (ii) retail price decreases in inventory.[12] In this section, we estimate the effect of inventory on price changes separately from other factors and find empirical facts to be consistent with our theoretic predictions.

Second, we document and decompose the retail price dispersion in the used-car market. Specifically, we tease out the contributions of obvious factors such as seasonality and heterogeneities

[10] An inventory management expert Jasen Rice of LotPop said "For a dealer having 50 units or fewer on the lot, one or two inventory management mistakes can crush their month." See https://www.cbtnews.com/dealers-experts-discuss-inventory-holding-cost-erosion/ for details.

[11] Our general understanding of the industry is from various industry reports, including Edmunds' "Used Vehicle Market Report," Manheim's "Used-Car Market Report," and Murry and Schneider (2015). For industry reports, see `https://dealers.edmunds.com/static/assets/articles/2017_Feb_Used_Market_Report.pdf` and `https://publish.manheim.com/content/dam/consulting/2017-Manheim-Used-Car-Market-Report.pdf`

[12] As illustrated in section 4, these implications are robust to realistic perturbations to accommodate various heterogeneities in the used-car market.

among intermediaries and products, and characterize the residual distributions of inventories and retail prices. We find that there exists a substantial dispersion among residual prices, and the distribution is unimodal. Recall that our model offers a novel prediction: in the absence of any ex-ante heterogeneity, the ex-post differential inventory holding can lead to a substantial price dispersion, and the distribution is unimodal. Therefore, our model provides a unified explanation of all these empirical facts.

5.1 Data and Descriptive Statistics

We obtain information on used-car listings from a large car listings platform, `cars.com`. We observe the daily listings for dealers who list inventory on the platform in the state of Ohio in 2017. For each car, we know the Vehicle Information Number (VIN), which is a unique number assigned to a vehicle that contains information to describe and identify the vehicle, make, model, model year, and trim with a particular set of options, exterior color, odometer mileage, whether it is certified by the OEM, and the daily listing price from the date when it is initially listed to the date when it is removed from the website.

Notably, the platform's pricing is not marginal to the number of cars listed, and the platform reports that dealers typically list their entire inventory on the platform. Moreover, according to our conversations with `cars.com`, most car dealers update their listings on the platform immediately. Therefore, we are confident that a dealer's new listings, listing removals, and active listings at a point of time are the actual new orders, car removals, and inventory in the dealership at that time, respectively. Although our data do not allow us to identify where a newly added car is obtained from and where a removed car goes to, a dealer's new orders and car removals at a point of time are good measures of the inflows and outflows of that dealer's inventory, which is our primary focus.[13]

Products. While our model analyzes homogeneous products, in reality, cars are highly differentiated. To better harmonize the data with the theoretical model, we focus on sedans and SUVs of non-luxury brands younger than 20 years old, and we group cars into different car types based on their body style and age. Later in our analysis, we treat each combination of these criteria as a distinct product category.[14] After selecting car types, we end up with 472,083 used cars, including 55% Sedans and 45% SUVs. We use four age categories: three years and younger (age group 1), four to six years (age group 2), seven to ten years (age group 3), and above ten years (age group 4). This dichotomization leaves us with eight product categories in total.[15]

[13] A newly added car can come from a wholesale trade-in or dealer-to-dealer auction market, or just be allocated from another site if the dealer is a chain store. Similarly, a removed car can be sold to an individual or another dealer or reallocated to another site if the dealer is a chain store.

[14] Non-luxury brands include Chevrolet, Chrysler, Dodge, Ford, GMC, Honda, Hyundai, Jeep, Kia, Mazda, Mercury, Mitsubishi, Nissan, Pontiac, Saturn, Subaru, Suzuki, Toyota, and Volkswagen. We exclude luxury cars in this study because it is too coarse to use body style and age to define the product of luxury cars, which are considerably heterogeneous even within the same brand and model. We examine the weekly new orders, car removals, and inventory separately for luxury brands, and the overall trends are similar despite their different levels.

[15] To be consistent with the model, we think of the categorizing the products from the dealers' perspective instead

Although a dealer typically sells multiple products, our model assumes that every intermediary sells a single product. To reconcile this, we treat a dealer and product category combination as a single decision-maker. Meanwhile, we control for the inventory holdings of all other cars by a dealer, including used cars and new cars, to capture the potential within-dealer inventory decisions. We also control for the inventory holdings of the same product by other nearby dealers to capture cross-dealer competitive concerns. Considering that dealers' portfolios of car inventories across product categories are quite stable over time, treating each product category as a decision maker is a reasonable assumption.

We only keep dealer-products that have positive inventory every week during 2017, similar to Aguirregabiria (1999). In the end, our product-dealer-level sample is a balanced panel of 3,121 product-dealers, 162,292 observations in the product-dealer-week level in total. Our car-level sample includes 2,026,679 car-week observations representing 340,277 cars listed by 743 dealers.[16]

Sample statistics. In the used-car market, dealers have heterogeneous capacity to hold inventory, so we expect that dealers' inventory processes are heterogeneous. Figure 1a reports the distribution of the total number of used cars in an average week, regardless of car types, held by the 743 dealers included in our sample.[17] Unsurprisingly, the average inventory holdings differ substantially across dealers. The majority of dealers hold fewer than 100 units on average, but some large dealers stock more than 300 units on average (for example, the national chain store CarMax). In Figure 1b, we report the distribution of the number of product categories (as defined above) that each dealer carries; recall that we balance our dealer sample so each dealer carries the same number of product categories each week.

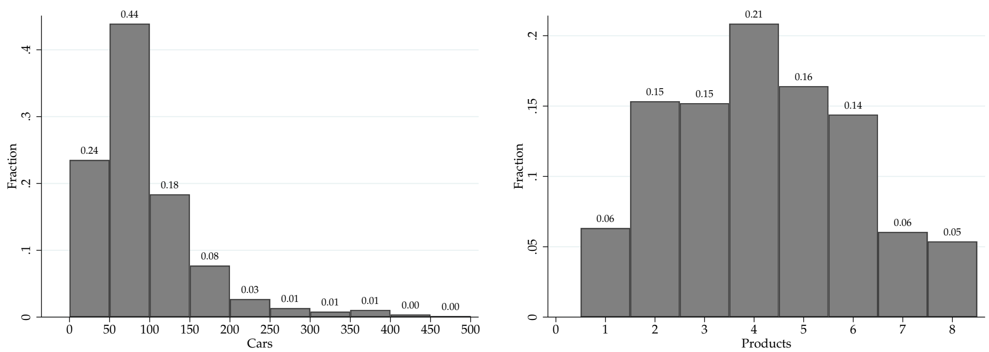
We present evidence that inventories and prices fluctuate *within* a dealer-product. Turning to Panel A of Table 1, the average (median) dealer stocks 13.05 (8.63) cars in a given week. The average (median) dealer's standard deviation of inventory, what we term inventory *volatility*, is 3.76 (2.80). The average dealer removes and adds 1.89 cars in a week. There is a lot of heterogeneity across dealers in inventories and inventory volatility. In Figure 2, we show the distribution of average dealer inventory holdings, and in Figure B.1 we display the distribution of inventory volatility across dealers.

In Panel B of Table 1, we present the descriptive statistics of all 340,277 listed cars in the sample, including the listing time, car ages, listing prices. We also calculate the mean and the standard deviation of each car's prices during its listing time and report the statistics in the last two columns in the panel. In the sample, 6.9% of vehicles were listed for only one week, and so we exclude

of that of consumers. This is because our reduced-form specification is meant to capture dealers' decisions rather than consumers' choices. We acknowledge that cars within the same product categories exhibit significant heterogeneity to individual consumers, but these individual preferences are also highly idiosyncratic. A dealer's preference for cars depends on the aggregate demand of consumers and the margins earned.

[16] If we considered these product-dealer combinations that contained zero inventories in our working sample, we would introduce a spurious upward bias in the frequency of zero orders and thus a downward bias in our estimate of the impact of inventory on orders. As a robustness check, we also repeat all empirical analyses using the full sample that includes those product-dealers with zero inventory at some time during our sample period. The sample statistics and results are reported in the appendix. Overall, the main results still hold.

[17] Notice that it includes all used cars, not only the eight products that we focus on but also all other used cars.



(a) Used-Car Inventory (b) Number of Product Categories

Figure 1: Dealer-Level Statistics: Average Inventory and Number of Products

Note: In Figure 1a, an observation is a dealer's total inventory of all used cars, including the eight products that we focus on and all other used cars as well, averaged over all weeks. In Figure 1b, an observation is number of products that a dealer holds inventory throughout the sample year. Both figures include 743 dealer-level observations in total.

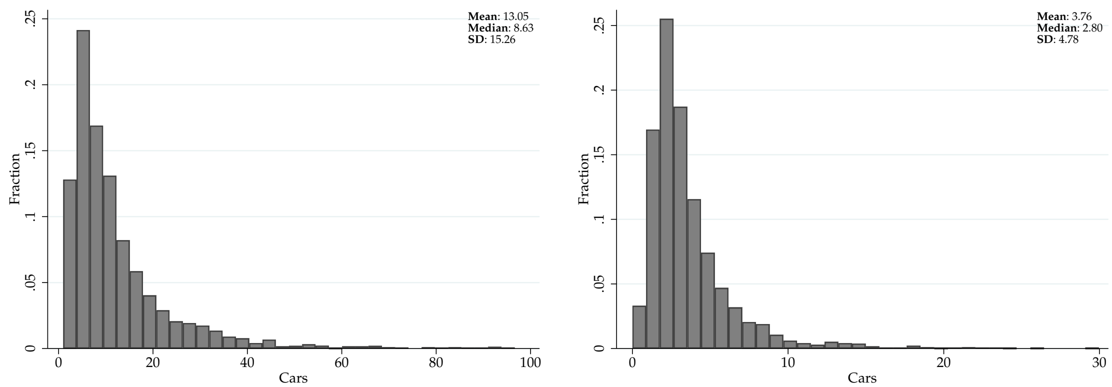
Table 1: Descriptive Statistics

Panel A. Product-dealer level†	Inventory		New orders		Car removals	
	Average	Volatility	Average	Volatility	Average	Volatility
Mean	13.05	3.76	1.89	1.63	1.89	1.46
Median	8.63	2.80	1.17	1.23	1.19	1.18
SD	15.26	4.78	3.62	2.30	3.60	2.08
Panel B. Car level‡						
			Listing prices ($)			
	Listing weeks	Car age	Initial price	Last price	Average	Volatility
Mean	7.83	4.93	16,407	15,702	16,121	376
Median	6	4	15,000	14,500	14,975	251
SD	7.15	3.18	8,040	7,795	7,927	508

Notes. Data source: Cars.com.

† An observation at the product-dealer level is the time-series average or standard deviation (volatility) of a product-dealer's inventory, new orders, and car removals over 52 weeks of 2017. The sample includes 3,121 product-dealer observations.

‡ Sample selection is described in text. The sample includes 340,277 car listings.



(a) Average Inventory

(b) Inventory Volatility

Figure 2: Heterogeneity in Dealer Inventories

Note: An observation is the time-series mean or SD of a product-dealer's inventory holdings over 52 weeks. Each figure includes 3,121 product-dealer observations.

those cars when we look at the car-level price volatility. The large dispersion of car ages and listing prices indicates car-level heterogeneity. Statistics of price volatility suggest that (i) price adjustments are common and significant and that (ii) such price adjustments differ substantially across cars. In fact, 57% of cars in the sample show price changes at least once, and among these cars, 20% experience price increases, and 5% have at least two price increases during their listing times.

5.2 Inventory and Transactions

First, we examine how dealers' decisions for how many new orders to place and how fast to sell relate to their inventory. To be specific, we run a regression of a dealer's new orders, or listing removals, in a week on that dealer's inventory at the beginning of that week, controlling for additional un-modeled factors that may be relevant in the data:

$$y_{kft} = \beta_0 + \beta_1 x_{kft} + \boldsymbol{Z}_{kft}\boldsymbol{\beta}_2 + d_{kf} + u_{kft}, \tag{17}$$

where y_{kft} is the log of dealer f's outcome (new orders or car removals) of product category k in week t, x_{kft} is the log of dealer f's inventory of product category k at the beginning of week t, $\boldsymbol{Z}_{kft}$ is a vector of variables to control for cross-product impacts within a dealer and cross-dealer impacts, d_{kf} is a full vector of product-dealer fixed effects, and u_{kft} is an idiosyncratic error term.

In our set of controls, $\boldsymbol{Z}$, we include the total inventory at the same dealer for other product categories to capture time-varying within-dealer financial and capacity related factors that may correlate with inventory. To control for time varying aggregate market trends, we include log of total inventory of the same product hold by all other dealers in the same zip code at the beginning of a week. We include a product-category and dealer dummy interaction to capture dealer-specific factors related to inventory management. For example, Figure 2 shows significant heterogeneity among dealers' inventory management decisions, suggesting the necessity of including product-dealer fixed effects.

Dynamic Panel IV Inventory, x_{kft}, is a function of last period's orders and sales through the following relationship,

$$x_{kft} = x_{kf,t-1} + orders_{kf,t-1} - sales_{kf,t-1} = x_{k,f,t-2} + orders_{k,f,t-2} \\ -sales_{kf,t-2} + orders_{kf,t-1} - sales_{kf,t-1} = ..., \tag{18}$$

so there is a lagged dependent variable in equation (17). To estimate the empirical specification, we take first-differences to eliminate the product-category-dealer interaction.

$$\Delta y_{kft} = \beta_1 \Delta x_{kft} + \Delta \mathbf{Z}_{kft} \boldsymbol{\beta}_2 + \Delta u_{kft}. \tag{19}$$

However, we expect OLS to be biased because inventory is a function of inventory last week plus orders and minus sales, one of which is the dependent variable. So $x_{k,f,t}$ depends on $u_{kf,t-1}$ and hence, Δx_{kft} is correlated with Δu_{kft}, that is, $Cov(\Delta x_{kft}, \Delta u_{kft}) \neq 0$.

To deal with this endogeneity problem, we use the inventories lagged at least $s^* + 1$ weeks as instrumental variables (IV) for the change of inventory Δx_{kft}, where $s^* \geq 1$. Current inventory is correlated with past inventory because the current inventory is the result of inventories, orders, and sales in the past. The validity of the instruments also relies on an exogeneity assumption: there exists a cutoff $s^* \geq 1$ such that $Cov(x_{kft}, u_{kf,t+s}) = 0, \forall s \geq s^*$. That is, the inventory at the beginning of a week is not correlated with the shocks occurring s^* weeks in the future and later. Under such an assumption, inventory at least $s^* + 1$ weeks before week t is not correlated with $\Delta u_{k,f,t}$. A potential threat is the possible serial correlation of the error term in the data. In our setting, the serial correlation may be caused by persistent demand or supply shocks. Forward-looking dealers may have rational expectation about those shocks and manage their inventory ahead of time. Consequently, past inventories are still correlated with the current shocks. However, the longer the lagged inventory is, the more confidence we have in our exclusion assumption. In our sample, three quarters of cars are on sale for shorter than two months. So it is reasonable to believe that dealers do not adjust their inventory in response to shocks two months in the future. In our following analysis, we will report the two-month lagged IV, but we have robustness to different lag length in Appendix X.[18]

[18] In Table B.1 in Appendix B we report the first-stage results, where the dependent variable is the change of log of the inventory at the beginning of a week, $\Delta x_{k,f,t}$, and the instruments are the log of inventory with various lags. The coefficient on lagged inventory are all negative and statistically significant at a 1% significance level for each lag time-frame, implying that the lagged inventories are negatively correlated with the change of current inventory. Moreover, as we lag the inventory by more weeks, the coefficient before the IV becomes smaller in magnitude, which is consistent with a decay in the time-dependence of inventories.

Estimation Results We report estimation results in Table 2. The dependent variable for the first two columns is weekly new listings. The dependent variable for the second two columns is weekly removed listings. Columns (1) and (3) report estimates from OLS, and columns (2) and (4) report estimates using lagged two-month inventory as an IV for current inventory.

Our main interest is the coefficient before the log of inventory, the first row. for new orders, the coefficient implies an elasticity of new orders to inventory of approximately -0.64. If inventory goes up by one percent new orders decrease by -0.64 percent. At the sample average inventory of 13.05 , one more unit of inventory at beginning of a week would result in 0.14 unit fewer new orders in that week. We also find an elasticity of listing removals to inventory of approximately 0.85. In other words, if inventory increases by one percent, listing removals increase by 0.85 percent. At the sample average, one more unit of inventory of a product at the beginning of a week would result in 0.19 unit more removals of that product in that week. Both of these findings are consistent with our model predictions, summarized in Proposition 2.

Table 2: Effect of Inventory on New Orders and Listing Removals

	New Orders		Listing Removals	
	(1) OLS	(2) IV	(3) OLS	(4) IV
Log Inventory	-0.0518***	-0.6486***	0.4318***	0.8512***
	(0.0043)	(0.0675)	(0.0038)	(0.0597)
... other cars at same dealer	-0.0003***	-0.0029***	-0.0004***	0.0017***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
... rivals in zipcode	0.0041	0.0245***	0.0003	-0.0132***
	(0.0041)	(0.0054)	(0.0037)	(0.0048)
... rivals in all zipcodes	0.0000***	0.0000***	0.0000***	-0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R^2				
First Stage F-stat				

Notes. The dependent variable is the log of new orders of a product placed by a dealer in a week. The sample includes 3,121 product-dealer observations and 162,292 product-dealer-week observations. Sample selection is described in text. Data source: Cars.com. Standard errors are in parentheses. $^{*}p < 0.10$. $^{**}p < 0.05$. ***

Inventory change has a larger impact in magnitude on the removing cars than on placing new orders. We interpret this as dealers relying more on the retail channel than the wholesale channel to manage their inventories. This is consistent with results in our quantitative analysis (see section 6.3).

5.3 Inventory and Retail Price

Next, we describe how prices are related to inventory levels and test the prediction of Proposition 2: retail prices decrease with inventory. We observe weekly prices for every car listed, so our unit of observation is a car-week.

To be specific, we run a regression of a cars' weekly listing prices on the inventory for that dealer and product-category (e.g. 4-6 year-old non-luxury sedan) at the beginning of that week, controlling for confounding factors,

$$p_{jkft} = \gamma_0 + \gamma_1 x_{kft} + \gamma_2 w_{jkft} + \mathbf{Z}_{kft}\gamma_3 + \mu_{jkf} + \epsilon_{jkft}, \tag{20}$$

where p_{jkft} is the log of the listing price of car j of product-category k listed by dealer f in week t, x_{kft} is the log of the dealer f's inventory of cars of the product k at the beginning of the week t, w_{jkft} is number of weeks that the car j has been listed at that dealer, $\mathbf{Z}_{k,f,t}$ is a vector of control variables which are the same as in equation (17), μ_{jkf} is a listing fixed effect, and ϵ_{jkft} is an error term. The listing effect implies that the co-variation in the data we use is within car-listing, e.g. changes in the price of same exact car listed at the same dealer over many weeks.

For estimation, we proceed analogously to the sales/acquisitions estimation. We take the first difference to eliminate the listing effects, then we use lagged inventory to instrument for current inventory. Table B.2 reports the first-stage results.

Estimation Results In Table 3 we report the estimation results, where column (1) reports the OLS result and column (2) reports the IV (2-month lagged inventory) results. Our main interest is the coefficient before the log of the inventory of the same product category by the same dealer. Our estimate is that the listing price of a car is 0.0397% lower if the inventory of the same type (product category) of cars is 1% higher. Recall that we include *listing* effects in the regression, so the variation used in the regression is over time for *the same exact car listing*. The sample mean of price is $16,121 and the sample mean of inventory is 13.05 at the car-week level. Our preferred estimate suggests that at the sample average, a car's price is $19 lower when the dealer has one more unit inventory of the same type of car. This is comparable to result from Zettelmeyer, Morton, and Silva-Risso (2006) for new-car markets.

5.4 Heterogeneity in Inventory Management

We further examine whether inventory has heterogeneous impacts on orders, sales, and price depending on the dealer size. To do this, we split the sample into dealers with above and below median inventory. The results show that the inventory has a stronger impact on new orders, car removals, and prices for large dealers than for small dealers, although the estimate of the interaction term is not significant for new orders. See Table B.3 in Appendix B for details. Large dealers are twice as elastic in the retail channel, with an sales-inventory elasticity of roughly 1.1, compared to 0.6 for small dealers. Large dealers are much more responsive with pricing. The inventory elasticity for small dealers is small and not statistically different from zero, whereas the price-inventory elasticity for large dealers is roughly -0.04, slightly more elastic than the main result for the pooled regression. A possible rationale is that large dealers are more likely to employ more advanced inventory-based order and pricing algorithms.

Table 3: Inventory and Retail Price

	(1) OLS	(2) IV
Log Inventory	0.0006***	-0.0397***
	(0.0002)	(0.0024)
... other cars at same dealer	0.0000***	-0.0001***
	(0.0000)	(0.0000)
... rivals in zipcode	0.0006***	0.0008***
	(0.0002)	(0.0002)
... rivals in all zipcodes	0.0000***	0.0000***
	(0.0000)	(0.0000)
Listing week	-0.0042	-0.0113***
	(0.0031)	(0.0026)
R^2		
First Stage F-stat		

Notes. The dependent variable is the log of the list price of a car at beginning of a week. The sample includes 340,277 cars and 2,026,679 car-week observations. Sample selection is described in text. Data source: Cars.com. Standard errors are in parentheses. $^{*}p < 0.10$. $^{**}p < 0.05$. $^{***}p < 0.01$.

5.5 Dispersion

We examine the distributions of inventory and retail price. The inventory process exhibits both idiosyncratic heterogeneity across dealers (Figure 2) and seasonality (Figure B.1). Both of these factors contribute to the inventory dispersion. To further understand the source of dispersion, we tease out seasonality and dealer heterogeneity by regressing inventory on explanatory variables. We interpret the residuals as the inventory of a homogeneous product-dealer-week and characterize the residual dispersion.

Specifically, we assume that the cutoff inventory level $x^*_{k,f,t}$, which is also the mode of the steady-state inventory distribution by Proposition 3, of dealer f who sells product k at time t can be represented as a summation of three independent factors:

$$x^*_{k,f,t} \equiv x^*_k + x^*_{k,t} + x^*_{k,f}. \tag{21}$$

In equation (21), x^*_k is a constant across all dealers who sell product k for each time period, $x^*_{k,t}$ captures the common seasonality shock shared by all product k dealers such that $\mathbb{E}_t(x^*_{k,t}) = 0$ where the expectation is taken across time t, and $x^*_{k,f}$ is the time-invariant fixed effect of dealer f such that $\mathbb{E}_f(x^*_{k,f})$ where the expectation is taken over dealers f. If $x^*_{k,t} = x^*_{k,f} = 0$ for every k, f and t, then all product k dealers are homogeneous and time-invariant, and the cutoff level of inventory is simply x^*_k for all product k dealers in every period. We further assume that the unconditional expectation of inventory of product k is x^*_k.

Empirically, let $x_{k,f,t}$ denote dealer f's inventory of product k at the beginning of time period t. In our application, a time period is one week. We construct a *normalized inventory* in a time period by double demeaning:

$$\tilde{x}_{k,f,t} = x_{k,f,t} - \bar{x}_{k,f} - \bar{x}_{k,t} + \bar{x}_k, \tag{22}$$

where $\bar{x}_{k,f}$ is the average of dealer f's inventory of product k over all weeks, $\bar{x}_{k,t}$ is the average inventory of product k in week t across all dealers, and $\bar{x}_k$ is the average inventory of product k across all dealers and over all weeks. Given our specification, it is straightforward to see that $\bar{x}_{k,f}$, $\bar{x}_{k,t}$ and $\bar{x}_k$ are consistent estimators of $x_k^* + x_{k,f}^*$, $x_k^* + x_{k,t}^*$ and x_k^*. Hence, the normalized inventory $\tilde{x}_{k,f,t}$ removes the impact of seasonality and idiosyncratic heterogeneity and estimates the difference between the actual inventory and the time-invariant common cutoff inventory level x_k^* for all product k dealers.

We test the unimodality of the inventory distribution by using the Dip Test (Hartigan and Hartigan 1985). First, we test the unimodality of the inventory distribution for *each* product-dealer panel. Each panel has 52 observations. Among all 3,121 product-dealer panels, we cannot reject uni-modality for 81% of the dealers.[19]

Next, we examine price dispersion. Notice that idiosyncratic difference among cars within the same type contributes to differences in their prices, so first, we need to net out these obvious factors to understand the composition of price dispersion. We run a regression of the log of a car's list price in a week on the weeks that it has been on sale at that time, the log of its mileage, and the make-model-model year fixed effects. We look at the distribution of price residual of each product-dealer combination, each with 52 observations. Among all 3,121 product-dealer groups, 86% are unimodal.[20]

6 Quantitative Analysis

In this section, we calibrate the model to answer two questions. First, how much residual list price dispersion can inventory management rationalize? Second, how much value does inventory management create?

Our calibration uses data of the most popular product, 4-6 years non-luxury sedans. This subsample includes 41,418 cars listed by 501 dealers, 223,902 car-week-level observations and 26,052 dealer-week-level observations. One should interpret the resulting model as a characterization of a *representative* dealer in the used-car market.[21]

[19]The property of single-modality is preserved after aggregation. In Appendix B, we plot (i) the distribution of the normalized inventory pooled over products, dealers and weeks, and (ii) the distribution of normalized inventory pooled over dealers and weeks for each products.

[20]The property of single-modality is preserved after aggregation. In Appendix B, we plot the distribution of the price residuals pooled over cars and weeks.

[21]For policy analysis, it is necessary to enrich our stylized model to capture substantial heterogeneity in the data. This is certainly beyond the scope of our paper.

6.1 Identification and Parameterization

We begin with functional-form assumptions. Search frictions on both the retail and the wholesale markets are summarized by a parameter-free urn-ball matching function, i.e., $\phi_r(\theta) = 1 - e^{-\theta}$, and $\phi_w(\lambda) = 1 - e^{-\lambda}$. Peters (2000) and Burdett, Shi, and Wright (2001) provide the micro foundation of an urn-ball matching function as a limit result of a finite directed search game as the number of traders goes to infinity. We specify the inventory cost as a linear function of inventory, i.e., $c(x) = cx$ with $c \geq 0$ being the parameter to be determined.

The time unit is set to be a week to match the data frequency. The weekly interest rate is set at $\rho = 5\%/52$ to match a 5% annual rate. The remaining parameters to be calibrated are a buyer's utility u, the opportunity costs for a buyer and a seller to search in corresponding markets κ_b and κ_s, and the marginal inventory cost c. The free-entry conditions in equations (2) and (3) and the optimization conditions in equations (7) and (8) are natural candidates as targeted moments. Our calibration procedure follows a few steps to address them.

The first step is to use the transition probability matrix of inventory to recover the unobserved market tightness sequences $\{\theta(x), \lambda(x)\}$. It resembles the standard approach in the labor search literature (see, e.g., Menzio and Shi (2011) and Guo (2018) where the state variable is workers' employment status). The evolution of the equilibrium inventory distribution in equation (10) corresponds to that of a continuous-time Markov chain. Therefore, the transition probability matrix in one time unit has a first-order approximation M such that each row x has at most three non-zero entries: $M(x, x-1) = \phi_r(\theta(x))$, $M(x, x+1) = \phi_w(\lambda(x))$, and $M(x, x) = 1 - \phi_r(\theta(x)) - \phi_w(\lambda(x))$. Then, for each normalized inventory level x, we equate $M(x, x-1)$ to the average fraction of sellers with fewer cars next week to recover a sequence of $\phi_r(\theta(x))$. The tightness sequence $\theta(x)$ can therefore be recovered following the urn-ball matching function assumption. The wholesale market tightness sequence $\lambda(x)$ is similarly obtained by equating $M(x, x+1)$ to the average fraction of sellers that increase in inventory levels. The recovered tightness sequences $\{\theta(x), \lambda(x)\}$ allow us to identify the remaining four parameters κ_b, κ_s, u, and c without using the information of wholesale prices. Indeed, as we have $\theta(x)$ and $\lambda(x)$ at multiple inventory levels, our model parameters are overidentified.

To mitigate the impact of dealer heterogeneity in capacity and market seasonality on inventory dynamics, we use normalized inventory, defined in section 5.5, when we recover the market tightness. Notice that the normalized inventory has a mean of zero by construction, so some normalized x's are *negative*. It is not an issue because the only role of x's in the calibration exercise is to serve as indices connecting recovered tightness with retail prices.

The second step is to get buyers and sellers' search opportunity costs κ_b and κ_s using prices in the data and recovered tightnesses. Combining the retail-market free-entry condition (2) and FOCs (7) and (8) yields

$$p(\theta(x+1)) = \kappa_s \frac{1}{\phi_w'(\lambda(x))} + \kappa_b \left(\frac{1}{\phi_r'(\theta(x+1))} - \frac{1}{\psi_r(\theta(x+1))} \right).$$

Table 4: Calibrated Parameter Values

Parameter	Value	Description and target
ρ	5%/52	Weekly interest rate to match a 5% annual rate
κ_b	3,732 ($)	A buyer's opportunity cost of search, jointly picked with κ_s using prices and weekly rates of listing removals and new orders
κ_s	7,527 ($)	A seller's opportunity cost of search, jointly picked with κ_b
u	16,270 ($)	Unit utility, based on retail-market free-entry condition (2) and FOC (7)
c	1.282 ($)	Marginal cost of inventory, to match data-implied value function

For each x, we use the *average* retail price of cars sold at inventory level $x+1$ on the left-hand side and calculate tightness-related terms on the right-hand side. We obtain κ_b and κ_s via OLS regression of $p(x+1)$ on the two tightness-related terms. While, in principle, we need the corresponding terms at two distinct values of x only to recover for these two parameters, we use eleven normalized inventory levels centered around zero, i.e., $x \in \{-5, -4, ..., 4, 5\}$, to make the estimation more robust. Importantly, the empirical price dispersion is not hardwired in the calibration process or the model-generated retail price distribution, as we exploit only the *average* prices at a subset of normalized inventory levels.

The third step is to get parameter u using recovered κ_b and κ_s. From equations (2), (7), and (8), we derive

$$u = p(\theta(x)) + \frac{\kappa_b}{\psi_r(\theta(x))} \quad \text{and} \quad u = \frac{\kappa_s}{\phi'_w(\lambda(x))} + \frac{\kappa_b}{\phi'_r(\theta(x+1))}.$$

Therefore, u is set to be the average of the evaluations of the two equations over all xs.

The last step is to compute $V(x+1) - V(x)$ and then c. Similar to the previous step, the FOCs (7) and (8) imply two ways to obtain $V(x+1) - V(x)$, such that

$$V(x+1) - V(x) = u - \frac{\kappa_b}{\phi'_r(\theta(x+1))} = \frac{\kappa_s}{\phi'_w(\lambda(x))}.$$

The average evaluation at each x generates a sequence of $V(x+1) - V(x)$. Taking a first-order difference of the dealer's value function in equation (6) yields the relationship between marginal cost c and $V(x+1) - V(x)$. We set c to be the average value implied by this relationship.

Table 4 reports the parameter values produced by the procedure described above. Most of the estimated parameters seem reasonable. Parameters κ_b and κ_s seem high but it should come with no surprise. Since we assume that buyers and sellers are short-lived, κ_b and κ_s reflect their *opportunity cost*, including both their physical search cost and (perhaps more importantly) their outside option values if they do not search for dealers. For instance, buyers and sellers can search for unmediated transactions. Under the free-entry specification, they also correspond to the *expected surplus* of buyers and sellers' market participation, adjusted by probabilistic matching. See our discussion of assumptions in the model section.

Our estimate of u captures the average monetary-measured utility of purchasing a 4-6-year-old non-luxury sedan. The marginal cost of inventory may seem smaller than expected. In principle,

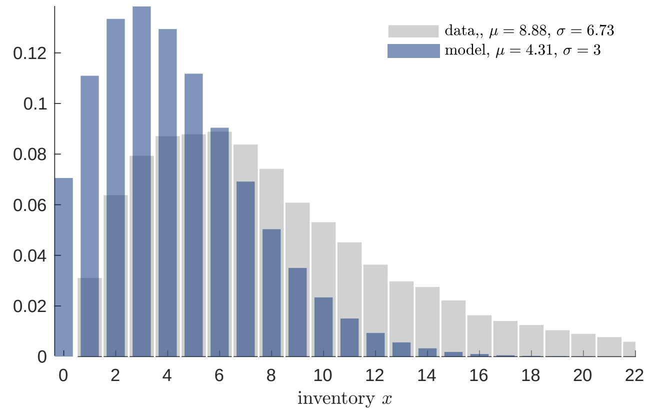


Figure 3: Inventory Distribution: Data and Model

Notes. Light-gray bars represent the frequency histogram of inventory distribution in the data subsample of 4-6 years non-luxury sedans. The range shown is from the 5th percentile to the 95th of inventory levels. Darker blue bars are model-generated probability mass function (PMF) of inventory.

both the change in the risk of stock-out and inventory cost affect equilibrium the inventory-based pricing dynamics. Our calibrated model attributes the dealer's trade-off of inventory management and dynamic pricing mainly to the former, highlighting the importance of market frictions. Additionally, it is important to consider the sources of marginal inventory costs. Likely the largest factor is the cost of debt, as many dealers use debt to finance inventory. At a 5% rate, the weekly debt payment for a car with a wholesale price of $10,000 is roughly $10, the same order of magnitude as our calibrated cost. Finally, in our calibration, the inventory cost parameter absorbs remaining dealer heterogeneity and other factors that affect dealers' flow revenue because we have no other parameter to reconcile marginal revenues with costs. Factors that capture benefits to dealers for keeping high inventories (e.g. marketing or search agglomeration), are included in our calibration of inventory costs.[22]

6.2 How much Price Dispersion can Inventory Management Explain?

We simulate the model using parameter values in Table 4 and compare the simulated statistics with their data counterparts. Although our model is overidentified, we do not target price or inventory distributions to calibrate the model. Therefore, it is worthwhile to compare the full working sample with model simulations.

Figure 3 plots the empirical and model-implied inventory distribution without normalization. Quantitatively, the model generates about half of the empirical mean and standard deviation. The asymmetry in the simulated distribution is due to the model restriction of $x \geq 0$ and $c \geq 0$.

Next, we examine the model's performance in replicating the empirical price distribution. The model-implied average retail price at the steady state is $\mathbb{E}[p^*(x) \mid \theta^*(x) > 0] = \$11,775$, close to the sample mean of raw data prices $11,753. This is not surprising because we used the average

[22]See Murry and Zhou (2020) for a discussion of search agglomeration economies among car dealers.

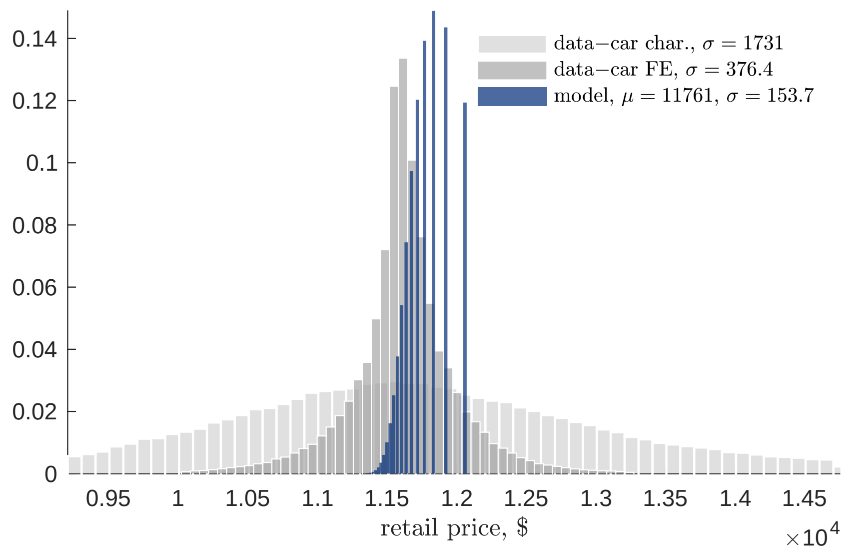


Figure 4: Price Distribution: Data and Model.

Notes. Gray bars show the empirical price distributions in the data subsample. Lighter-gray bars represent the frequency histogram of the empirical distribution of residual prices after controlling for observable idiosyncratic characteristics (e.g., car make, model-year, mileage, etc.) and time-on-market effects. Darker gray bars represent the frequency histogram of the distribution of residual prices after controlling for individual car fixed effects and time-on-market effects. The price range shown is from the 5th percentile to the 95th of residual prices. Dark blue bars are model-generated probability mass function (PMF) of retail prices.

raw price to calibrate the model. The model generates a \$154 standard deviation in retail price, which accounts for 4.33% of the standard deviation \$3,549 in the *raw* price data. However, our raw data include price differences across different makes-models of cars, something that our model doesn't capture. So instead, we consider dispersion in residualized prices. First, we residualize prices by regressing prices on car characteristics, including car make, model-year, and mileage, as in section 5.5. The residual dispersion remains large as cars differ in more aspects than we control, but the model-generated standard deviation is 8.88% of the residuals' \$1,731 and the 90-10 inter-percentile range is 12.0% of the residuals' \$4,052. We interpret these results as the lower bound of the model's explanatory powers, as such a homogenizing procedure removes only a portion of car-level heterogeneity and does *not* remove any dealer-level heterogeneity. Second, we residualize prices by additionally controlling for individual car fixed effects and time-on-market effects. The model-generated standard deviation is 40.8% of the price residuals' (\$376). The model distribution's 90-10 inter-percentile range is 57.6% of the residuals' (\$841). We view these results as the model's maximal explanatory power as the second procedure removes part of dealer-level heterogeneity which may include inventory variation. The reason is that each car is exclusively sold by one dealer, and removing car-level fixed effects partially controls for dealer characteristics.

We visualize the model-generated price dispersion and the empirical residual price dispersion in Figure 4. The shape and location of the simulated price distribution are governed by both the inventory distribution and the price policy as a function of inventory levels. The light grey bars in 4 represent the distribution of prices when we residualize prices by make/model/year. The darker grey bar represents price residuals after controlling for individual car effects, and the blue bars represent the model predicted price distribution.

We do not observe wholesale prices in the data. Nonetheless, the parameterized model implies an average wholesale price of $\mathbb{E}[w^*(x) \mid \lambda^*(x) > 0] = \$8,920$ at the steady state. The average wholesale and retail prices in turn imply that an average dealer charges buyers a 32% markup.[23] The implied markup level is consistent with Murry and Schneider (2015) in the literature.

It is useful to know what to expect when introducing heterogeneity among products and dealers to our stylized model. In Appendix C, we offer preliminary insights by considering the effect of different inventory and search opportunity costs (dealer heterogeneity in storage technology and prominence) and idiosyncratic match between buyers and used-cars (product differentiation).

6.3 The Value of Inventory Management

This section uses the calibrated model to assess the value of inventory management policy. In our model, dealers make trade possible. To participate in this mediated market, buyers and sellers' opportunity costs are κ_b and κ_s, respectively. Since buyers and sellers break even in the equilibrium, a dealer reaps all the surplus it creates. Therefore, $V(x)$ of a dealer with inventory x corresponds to the total continuation surplus he will create. Based on the calibrated parameters, at the steady-state, a dealer's ergodic surplus creation is $\mathbb{E}[V(x)] = \$856,815$. This value corresponds to the long-run surplus creation of mitigating search costs from a representative dealer selling a single product category. This number is intuitive, as it represents the discounted value of the markup (roughly \$2,750) multiplied by the weekly trades (0.29 chance of a trade each week).

In our model, the value created by dealers comes from two sources. One is to facilitate trading, and the other is to actively manage its inventory levels via inventory-based pricing and ordering. Therefore, we decompose this value into two corresponding parts and disentangle the two sources. The first part of social value is due to the mere existence of dealers. The remaining part thus reflects the pure value of inventory management.

6.3.1 Value of Intermediation

To compute the value of dealers' existence, we consider a dealer who can hold at most one unit of inventory. When it has no inventory, it can only order but not sell. When it has one unit of inventory, it can only sell but not order. This dealer's Bellman equation can be written as follows

$$rV^1(1) = -c + \max_{\theta} \phi_r(\theta)[u + V^1(0) - V^1(1)] - \kappa_b \theta,$$
$$rV^1(0) = \max_{\lambda} \phi_w(\lambda)[V^1(1) - V^1(0)] - \kappa_s \lambda.$$

[23] Due to inventory management and hence various retail and wholesale prices at different x's, we cannot define the markup at each inventory level.

At the calibrated parameter values, the ex-ante optimal value of the dealer is $\mathbb{E}^1[V^1(x)] = \$460,045$, where the expectation operator $\mathbb{E}^1$ is defined by the stationary distribution over $x \in \{0,1\}$. We interpret the difference between $\mathbb{E}^1[V^1(x)]$ and the baseline $\mathbb{E}[V(x)]$ as the value of holding and managing multi-unit inventory. The difference is a substantial (roughly 46.31%) part of the dealer's total value).

Another way to view this result is that multi-unit inventory dealers mitigate the risk of a stock-out, therefore increasing the likelihood of a sale at any given moment. We attribute $\mathbb{E}[V(x)] - \mathbb{E}^1[V^1(x)]$ as the value of engaging in optimal inventory management that decreases the risk of stock-outs. This value would be small when the dealer finds it optimal not to hold fewer units of inventory than the baseline model predicts. This is the case when inventory and search costs are sufficiently large (c and κ), or the gain from trade is small. Notice that the marginal value of additional inventory capacity is decreasing. This is a natural consequence of the standard assumption that the matching technology has diminishing returns.[24]

6.3.2 Value of Inventory-Based Price Policies

Next, we take a closer look at the value of inventory management. Recall that a dealer trades off between the risk of stocking out and the cost of ordering/holding more inventory. An inventory-based pricing and ordering policy allows dealers to dynamically control the speed of inventory inflows relative to the speed of outflows. The dealer has two instruments to do so: the retail price and the wholesale order policies.

Freeze retail price policy. First, we fix a retail policy $\overline{\theta}$ as long as $x > 0$ and allow the dealer to optimally choose the wholesale price policy, λ. Therefore, the retail price $\overline{p}$ is constant over $x > 0$ according to equation (4). In this case, the value function at the given $\overline{\theta}$ is

$$\begin{aligned} rV^w(x,\overline{\theta}) = &- c(x) + \phi_r(\overline{\theta})[u + V^w(x-1,\overline{\theta}) - V^w(x,\overline{\theta})] - \kappa_b\overline{\theta} \\ &+ \max_{\lambda}\left\{\phi_w(\lambda)[V^w(x,\overline{\theta}) - V^w(x+1,\overline{\theta})] - \kappa_s\lambda\right\}. \end{aligned}$$

We consider the expected loss in surplus creation when a random dealer drawn from the baseline steady-state distribution is forced to "freeze" retail policy henceforth. If the randomly picked dealer has inventory x_0, then he must use the constant retail policy $\theta^*(x_0)$ from now on regardless of realized changes in inventory, but he is allowed to optimally choose his wholesale policy. Using the baseline steady-state distribution, we can calculate the expectation of the difference $\mathbb{E}_x[V(x) - V^w(x,\theta^*(x))]$. This value gap, capturing the expected loss in surplus due to losing control over inventory outflows,is $\$85,607$, which is about 10% of the total baseline surplus.

[24] By no means, our decomposition exercise implies that a dealer will benefit from splitting his business into multiple small scale entities. To mitigate the diminishing returns of the matching technology, these small scale entities must be located in different locations to attract different set of traders. An obvious hurdle to do so is the initial setup cost for each entity. In addition, splitting the business will probably alter the inventory cost.

Freeze wholesale price policy. Second, we instead fix the wholesale policy at $\overline{\lambda}$ and let the dealer optimally choose the inventory-based retail policy. Then the value function is

$$\begin{aligned} rV^r(x,\overline{\lambda}) = &- c(x) + \phi_w(\overline{\lambda})[V^r(x,\overline{\lambda}) - V^r(x+1,\overline{\lambda})] - \kappa_s\overline{\lambda} \\ &+ \max_{\theta}\left\{\phi_r(\theta)[u + V^r(x-1,\overline{\lambda}) - V^r(x,\overline{\lambda})] - \kappa_b\theta\right\}. \end{aligned}$$

Again, we consider the expected loss in surplus creation when a random dealer drawn from the baseline steady-state distribution is forced to "freeze" wholesale policy henceforth. The expected loss $\mathbb{E}_x[V(x) - V^r(x,\lambda^*(x))]$ is \$13,285, which is about 1.55% fraction of the total surplus.

Freeze retail and wholesale. A tempting next step is to force the dealer to set static policies for both retail and wholesale prices, further decomposing the value of holding and managing multi-unit inventory into two parts — the value of unconstrained inventory capacity and the pure value of optimally managing the inventory level. Unfortunately, it is not a straightforward task. Ideally, the value of inventory capacity would be the ex-ante social surplus that a dealer with unconstrained capacity creates using a fixed retail policy (θ) and a fixed wholesale policy (λ). The value of inventory management then follows. However, for an arbitrary pair of (θ,λ), the corresponding stochastic process of inventory $\{x_t\}$ may not be stationary. Since the inventory cost function is also unbounded, it is impossible to compute the dealer's value associate with such a policy and the corresponding ergodic inventory distribution. One practical solution is to introduce an exit option to the dealer. Once its continuation value $V(x)$ turns negative, the dealer exists the market and takes the outside option 0. This will lead to a well-defined value function $V^+(x,\theta,\lambda)$. Note that for a generic pair of (θ,λ), the dealer will exit almost surely after a sufficiently long time period, so there is neither a well-defined stationary inventory distribution nor a long-run value of the dealer.

Again, we consider the expected loss in surplus creation when a random dealer drawn from the baseline steady-state distribution is forced to "freeze" its policies henceforth with an option to exit. If the randomly picked dealer has inventory x_0, then he must use the same retail policy $\theta^*(x_0)$ and the same wholesale policy $\lambda^*(x_0)$ until exit, regardless of realized changes in inventory. The dealer's continuation value is $V^+(x_0,\theta^*(x_0),\lambda^*(x_0))$. Therefore, using the baseline steady-state distribution, we can calculate the expectation of the difference $\mathbb{E}[V(x) - V^+(x,\theta^*(x),\lambda^*(x))]$. This value gap captures the expected loss in surplus due to losing control over inventory inflows and outflows, which is \$256,631, a large (29.95%) fraction of the total surplus.

Summary. We summarize the results in Table 5. The existence of dealers (i.e., the ability to hold at most one unit of inventory) accounts for slightly over one half of the total social surplus they create, and the rest is due to two-sided inventory management. Inventory management affects retail and wholesale prices, but what matters the most to the creation of social surplus is the changes in trading frequencies when prices shift. A binding inventory capacity constraint greatly restricts dealers' control over trading speed. In contrast, even if dealers have limited options

Table 5: Value of holding and managing inventory

	Social surplus	Prob. of trade	Retail price	Wholesale price
Baseline model	$ 856,815	0.29	$ 11,775	$ 8,920
	% Baseline			
Fixed capacity $x \in \{0, 1\}$	53.69	55.17	98.75	98.74
Frozen retail†	90.01	90.46	–	99.44
Frozen wholesale‡	98.45	98.95	100.10	–
Frozen retail and wholesale	70.05	–	–	–

Notes. Baseline model results are reported as raw simulated values. Other results are shown as percentage points of Baseline. Social surplus is the corresponding $\mathbb{E}[V(x)]$, $\mathbb{E}^1[V^1(x)]$, $\mathbb{E}[V^r(x, \theta^*(x))]$, or $\mathbb{E}[V^w(x, \lambda^*(x))]$, probability of trade is $\mathbb{E}[\phi_r(\theta^*(x))] = \mathbb{E}[\phi_w(\lambda^*(x))]$ at the baseline steady state, retail price is $\mathbb{E}[p^*(x) \mid \theta^*(x) > 0]$, and wholesale price is $\mathbb{E}[w^*(x) \mid \lambda^*(x) > 0]$.
† The trade probability is based on the mean of wholesale trade probability at the time of freezing the retail policy. The expected wholesale price is also calculated at the time of freezing.
‡ The trade probability is based on the mean of retail trade probability at the time of freezing the wholesale policy. The expected retail price is also calculated at the time of freezing.

and can only actively manage inventory levels from one side, retail or wholesale, it is possible to achieve the majority of the benefit.

An interesting observation is that retail-side managing is more valuable than wholesale-side managing. It is a quantitative result due to (i) a higher marginal social cost on the wholesale side κ_s than on the retail side κ_b, and (ii) a large calibrated utility u on the retail side. Adjusting the wholesale strategy λ is associated with a larger marginal cost but a smaller marginal benefit relative to the retail strategy, implying that the optimal wholesale policy is less sensitive to inventory changes; and thus has a smaller variation. As a result, freezing the wholesale policy has a smaller impact than freezing the retail policy. This result is consistent with our empirical findings in Table 2. The same logic further suggests that the retail price is more sensitive to inventory changes than the wholesale price, implying a larger dispersion in retail prices than that in wholesale prices. This is confirmed by the simulated model, where the standard deviation in the wholesale price is $150, which is less than that in retail prices.

7 Concluding Remarks

This paper fills a gap between several active areas of literature: one on search theoretic models of intermediaries, one on price dispersion, and one on pricing and inventory control. We highlight the role of inventory dynamics in shaping retail price dispersion and its dynamics in a search model. The natural combination of equilibrium search and inventory management has a significant logical consequence. We rationalize unimodal price distributions using a model where price dispersion results from intermediaries' optimal inventory and revenue management without relying on any ex-ante heterogeneity. Prices fluctuate as the intermediate finds itself away from the optimal inventory size, and intermediaries adjust prices to sell inventory or restock.

In our empirical application, we document various findings supporting this mechanism of price dispersion by analyzing a dataset on used-car listings. We calibrate the model's parameters and demonstrate that our proposed inventory-management channel generates a substantial price dispersion and the practice of inventory management significantly contributes to the surplus creation of intermediaries. Our stylized model is built to demonstrate the interplay between inventory management and search frictions in a transparent manner and to provide a tractable framework for future studies about intermediaries. It is unrealistic to expect a seamless fit to the rich data without adding significant complications that undermine the transparency of the current model. Our partial equilibrium model also ignores the general equilibrium effect that the existence of dealers will alter buyers and sellers' decision to enter unmediated markets and change matching efficiency (Gavazza 2016 and Salz 2017). For further welfare analysis and policy evaluation, it would be important to model (i) heterogeneity among dealers, consumers, sellers, and products, and (ii) the above-mentioned cross-market search externalities.[25] We leave these issues for future work.

Finally, while our empirical application has focused on used-car markets, the insights from the model are meant to capture the economics of inventory management more broadly.

Acknowledgements We thank Victor Aguirregabiria, Simon Anderson, Gary Biglaiser, Michael Choi, Luca Flabbi, Chao Fu, Qing Gong, Michael Grubb, Naijia Guo, Veronica Guerrieri, Guido Menzio, Giuseppe Moscarini, Peter Norman, Stan Rabinovich, Paulo Somaini, Andrei Shevchenko, David Wiczer, Randall Wright, Shouyong Shi, Valentin Verdier, Jidong Zhou, Qiankun Zhou, participants of 2019 Mini conference on search and money at Madison, 2019 Midwest Macroeconomic conference at East Lansing, 2020 SaMMF Workshop: Advances in Search Theory, 2021 IIOC, Stanford IO Seminar, and the University of Virginia Theory/IO seminar for helpful comments.

A Appendix: Proofs

Proof of Lemma 1. Suppose that an intermediary holds x units of inventory at time t and it employs the following policy $\Gamma \equiv \{\theta_\tau, \lambda_\tau\}_{\tau \geq t}$ that generates an inventory process $\{x_\tau\}_{\tau \geq t}$. The policy solves the problem of (1) for the inventory being $x_\tau + 1$ until the first instant when its true inventory inventory drops to $x - 1$. Denote

$$T \equiv \inf\{\tau \geq t : x_\tau \leq x - 1\}.$$

For $\tau \geq T$, it employs the optimal policy. Denote the associate life-time profit to be $V^\Gamma(x)$, then we must have

$$V^\Gamma(x) = V(x+1) + \mathbb{E}\left\{\int_t^T e^{-\rho(\tau-t)}[c(x_\tau+1) - c(x_\tau)]d\tau + e^{-\rho(T-t)}[V(x-1) - V(x)]\right\}$$

where the expectation is taken over the random time T. That is, $V^\Gamma(x)$ differs from $V(x+1)$ in aspects. First, in time interval $[t, T)$, the flow inventory cost is $c(x_\tau)$ rather than $c(x_\tau + 1)$. Second,

[25] As we discussed in Section 2, the general equilibrium effect can be accommodated when the measure of buyers and sellers is fixed in each period.

after time T, the continuation value is $V(x-1)$ instead of $V(x)$. Because policy Γ is suboptimal, $\hat{V}(x) \leq V(x)$; and therefore

$$V(x+1) - V(x) \leq \mathbb{E}\left\{ -\int_t^T e^{-\rho(\tau-t)}[c(x_\tau+1) - c(x_\tau)]d\tau + e^{-\rho(T-t)}[V(x) - V(x-1)] \right\}. \quad (23)$$

Moreover, because $c(\cdot)$ is increasing and $\mathbb{E}[e^{-\rho(T-t)}] \in (0,1]$, the right-hand side of (23) is less than or equal to $V(x) - V(x-1)$. As a consequence, $2V(x) \geq V(x-1) + V(x+1)$ for $x = 1, 2, ...$, or $V(x) - V(x-1)$ decreases. □

Proof of Proposition 2. Recall that both $\phi_r(\cdot)$ and $\phi_w(\cdot)$ are increasing. From Lemma 1, both $V(x) - V(x-1)$ and $V(x+1) - V(x)$ in FOCs (7) and (8) are decreasing in x, so the first part of the proposition immediately follows. The second part of the proposition is a direct consequence of the combination of part 1 and conditions (4) and (5). □

Proof of Proposition 3. First, we prove the existence and uniqueness. In equilibrium, a dealer's inventory follows an ergodic finite state Markov process $\{x_t\}$ determined by the equilibrium policy $\theta^*(\cdot)$ and $\lambda^*(\cdot)$. The Markov process is asymptotically stationary and has a unique invariant distribution, which satisfies (11) for each $x = 0, 1, \ldots, s$ where $s \in \mathbb{N}$ is the base level of the stock defined in (9).

Second, we prove that the steady-state distribution is unimodal. Rearranging (11) yields

$$\phi_w(\lambda^*(x))g_{ss}(x) - \phi_w(\lambda^*(x-1))g_{ss}(x-1) = \phi_r(\theta^*(x+1))g_{ss}(x+1) - \phi_r(\theta^*(x))g_{ss}(x). \quad (24)$$

Because $\phi_w(\lambda^*(x))$ decreases in x, the left-hand side of (24) is less than $[g_{ss}(x) - g_{ss}(x-1)]\phi_w(\lambda^*(x-1))$. Because $\phi_r(\theta^*(x))$ increases in x, the right-hand side of (24) is greater than $[g_{ss}(x+1) - g_{ss}(x)]\phi_r(\theta^*(x))$. Therefore, we have $[g_{ss}(x) - g_{ss}(x-1)]\phi_w(\lambda^*(x-1)) \geq [g_{ss}(x+1) - g_{ss}(x)]\phi_r(\theta^*(x))$. That is, for any $x \geq 1$, whenever $g_{ss}(x+1) \geq g_{ss}(x)$, we have $g_{ss}(x) \geq g_{ss}(x-1)$, and whenever $g_{ss}(x) \leq g_{ss}(x-1)$, we have $g_{ss}(x+1) \leq g_{ss}(x)$. So the steady-state probability mass function $g_{ss}(\cdot)$ is single-peaked, or unimodal.

□

B Additional Figures and Tables

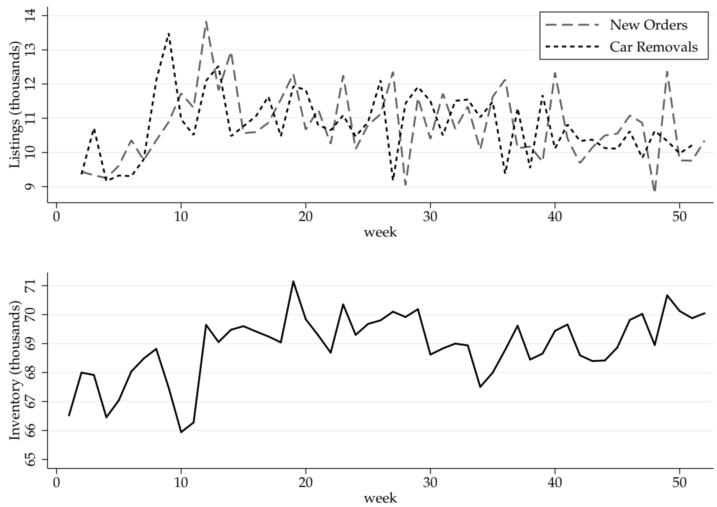


Figure B.1: Used-Car Inventory, New Orders, and Car Removals (Thousands)

Note: The dash, short dash, and solid lines depict the number of total new orders, the number of car removals, and the number of inventories of Ohio car dealers in every week from January 2017 to December 2017, 52 weeks in total. Data source: Cars.com.

Table B.1: First-Stage Results: The New Orders and Sales Equation (19)

	(1)	(2)	(3)	(4)
Log of inventory two weeks ago	-0.0222*** (0.0006)			
Log of inventory four weeks ago		-0.0147*** (0.0006)		
Log of inventory six weeks ago			-0.0107*** (0.0006)	
Log of inventory two months ago				-0.0067*** (0.0007)
Δ Log of inventory of other cars by same dealer	-0.0037*** (0.0000)	-0.0045*** (0.0000)	- 0.0045*** (0.0000)	-0.0044*** (0.0000)
Δ Log of inventory of same product by other dealers of same zipcode	0.0061*** (0.0028)	0.0065*** (0.0028)	0.0067*** (0.0029)	0.0057*** (0.0030)
Δ Log of inventory of same product by dealers of different zipcodes	0.0000*** (0.0000)	0.0000*** (0.0000)	0.0000*** (0.0000)	0.0000** (0.0000)
R square	0.4888	0.2704	0.1671	0.1087

Note: The dependent variable is the change of the log of the inventory at the product-dealer-week level. The sample includes 3,121 product-dealer observations and 162,292 product-dealer-week observations. Sample selection is described in text. Data source: Cars.com. Standard errors are in parentheses. $^{*}p < 0.10$. $^{**}p < 0.05$. $^{***}p < 0.01$.

Table B.2: First-Stage Results: The Price Equation (**??**)

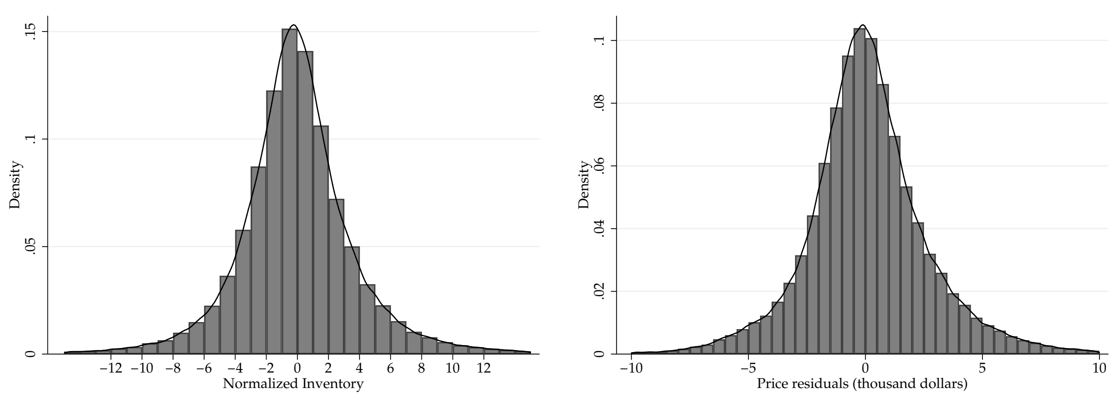
	(1)	(2)	(3)	(4)
Log of inventory two weeks ago	-0.0220*** (0.0001)			
Log of inventory four weeks ago		-0.0191*** (0.0001)		
Log of inventory six weeks ago			-0.0161*** (0.0001)	
Log of inventory two months ago				-0.0135*** (0.0001)
Δ Log of inventory of other cars by same dealer	-0.0019*** (0.0000)	-0.0023*** (0.0000)	-0.0023*** (0.0000)	-0.0023*** (0.0000)
Δ Log of inventory of same product by other dealers of same zipcode	0.0084*** (0.0007)	0.0091*** (0.0007)	0.0096*** (0.0008)	0.0097*** (0.0008)
Δ Log of inventory of same product by dealers of different zipcodes	0.0000*** (0.00000)	0.0000*** (0.00000)	0.0000*** (0.00000)	0.0000** (0.00000)
R square	0.1394	0.1462	0.1452	0.1418

Note: The dependent variable is the change of the log of the inventory at the car-week level. The sample includes 340,277 cars and 2,026,679 car-week observations. Sample selection is described in text. Data source: Cars.com. Standard errors are in parentheses. $^{*}p < 0.10$. $^{**}p < 0.05$. $^{***}p < 0.01$.

Table B.3: Heterogeneous Effects of Inventory: Dealer Size

	(1) Log(Orders)	(2) Log(Removals)	(3) Log(Price)
Log of inventory of same product by same dealer	-0.5668***	0.5938***	-0.0031
	(0.0712)	(0.0639)	(0.0059)
*Large Dealer	-0.1637	0.5153***	-0.0389***
	(0.1231)	(0.1104)	(0.0066)
Weeks on sale			-0.0107***
			(0.0027)
*Large Dealer			-0.0015***
			(0.0001)
Log of inventory of other cars by same dealer	-0.0032***	0.0026***	-0.0002***
	(0.0005)	(0.0004)	(0.0000)
Log of inventory of same product by other dealers of same zipcode	0.0265***	-0.0196***	0.0013***
	(0.0059)	(0.0053)	(0.0002)
Log of inventory of same product by dealers of different zipcodes	0.0000***	-0.0000***	0.0000***
	(0.00000)	(0.0000)	(0.0000)
R square	0.2710	0.4101	0.2931

Note: In all specifications, we use the log of the inventory lagged by two months as the instrument for the log of the current inventory. The sample used in the first two columns includes 3,121 product-dealer observations and 162,292 product-dealer-week observations. The sample used in the last column includes 340,277 cars and 2,026,679 car-week observations. Sample selection is described in text. Data source: Cars.com. Standard errors are in parentheses. $^{*}p < 0.10$. $^{**}p < 0.05$. $^{***}p < 0.01$.



(a) Inventory

(b) Price Residuals

Figure B.2: Distributions of Normalized Inventory and Price Residuals

Note: In left panel, an observation is a dealer's normalized inventory of a product at the beginning of a week, $\tilde{x}_{kft}$ defined in text. It includes 162,292 product-dealer-week observations in total. The black solid line is the kernel fitting of the distribution. In right panel, an observation is the price residual from a regression of the log of a car's list price in a week on time on market, log of mileage, and make-model-model year fixed effects. It includes 2,026,679 car-week observations in total. The black solid line is the kernel fitting of the distribution.

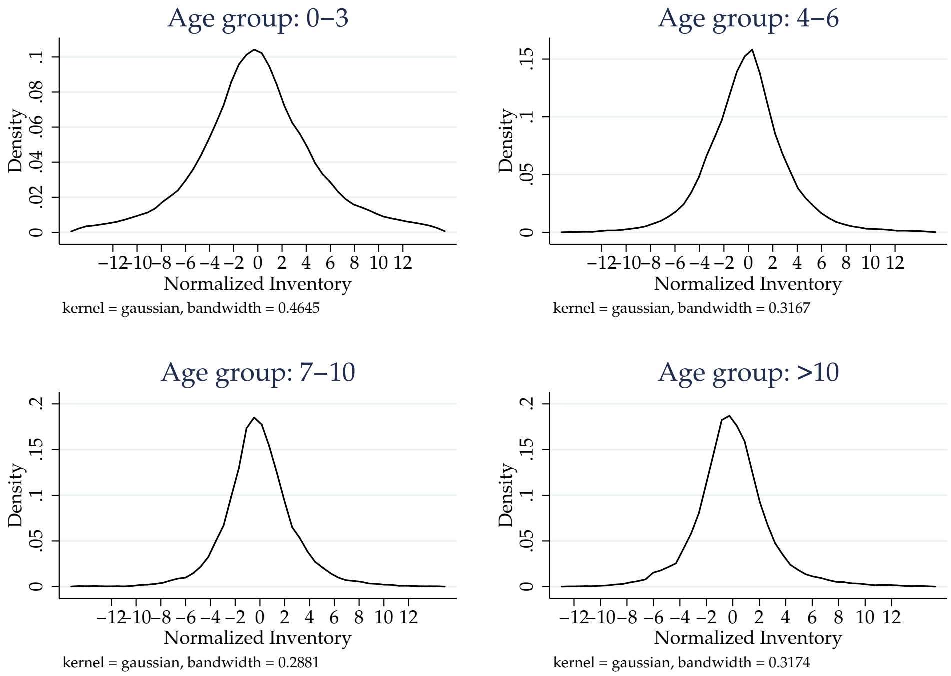


Figure B.3: Normalized Inventory of Sedan by Age Group

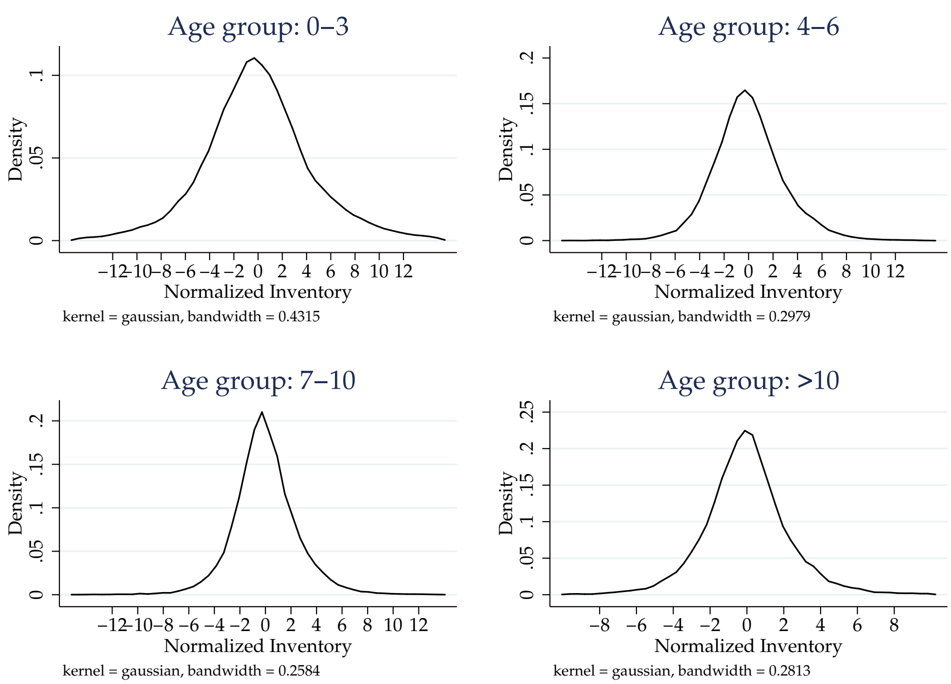


Figure B.4: Normalized Inventory of SUV by Age Group

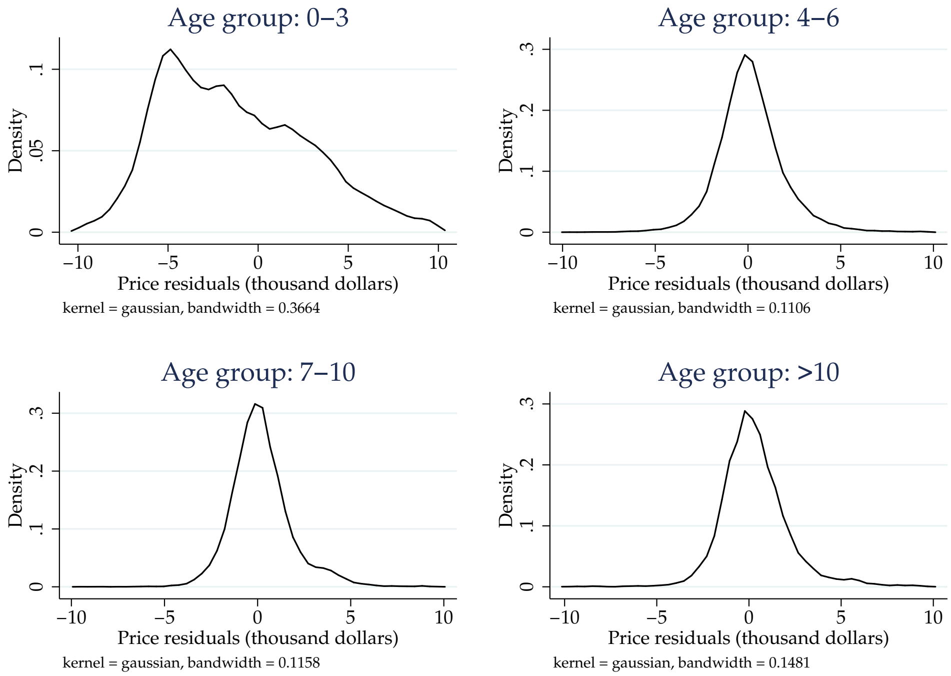


Figure B.5: Price Residuals of Sedan by Age Group

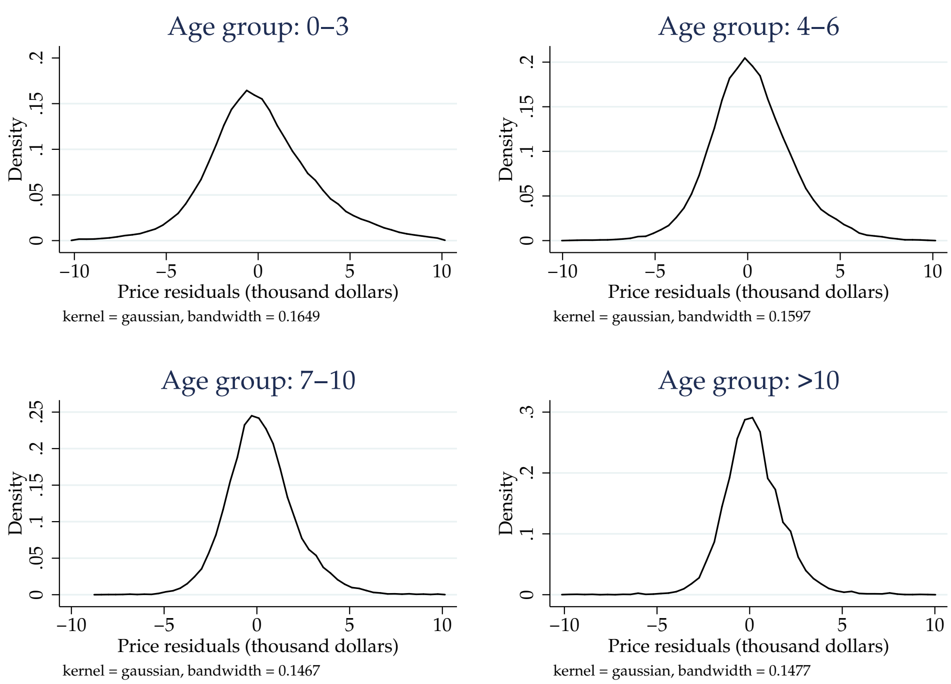


Figure B.6: Price Residuals of SUV by Age Group

C Further Quantitative Explorations

As discussed in Section 3.1, the main driving force of our proposed mechanism is search frictions on both retail and wholesale markets, rather than inventory costs. Indeed, the calibrated value of marginal inventory cost c is much smaller than utility u and search costs κ_b and κ_s. To further explore the relative importance of inventory costs and search frictions, we examine the changes in simulated inventory and price distributions when we alter the parameter values.

Specifically, this section aims to explore three things. The first is to examine the limited role of the conventional inventory cost in the model. The second task is to differentiate the impacts of two search costs κ_b and κ_s on the equilibrium outcome by separately altering their values.[26] This is unique to our model with search frictions on both retail and wholesale sides. Third, we introduce the possibility of mismatches into the model and examine a potential cause of multiple modes in price distributions.

Figure C.7 shows how the steady-state distribution of inventory or prices responds to alternative parameter values of c, κ_b, and κ_s, one parameter at a time. Figure C.7a illustrates that the marginal inventory cost c is not crucial for the model mechanism. The results under $c = 0$ differ little from the benchmark, and hence shutting down the conventional incentives for inventory management has little impact on the model, qualitatively or quantitatively. Increasing c results in a more concentrated inventory distribution with a smaller mode but a more dispersed price distribution; and reducing c has the opposite effect. The mechanism remains intact even if additional inventory brings additional benefits ($c < 0$). Indeed, if moments of inventory distribution are included as calibration targets without imposing $c \geq 0$, its value can become mildly negative.

The rest of Figure C.7 compares the impacts of changes in κ_b and κ_s. Recall that parameters κ_b and κ_s reflect both the physical search cost and outside option values (e.g., through unmediated trades) of buyers and sellers. Therefore, we believe the presence of online direct transaction platforms such as Craigslist.org helps to increase the value of κ_s and κ_b. As will be clear soon, the changes in κ_b and κ_s have a profound impact on the retail prices. Figure C.7b demonstrates the role of buyers' search cost κ_b. As expected, a larger (smaller) κ_b shifts the entire price distribution's location to the left (right). The shape of the price distribution remains largely intact. The equilibrium inventory distribution responds in the same direction and less in scale. Figure C.7c tells a different story about sellers' search cost κ_s. In contrast to the impact of κ_b, a larger (smaller) κ_s shifts the entire price distribution to the right (left), in the opposite direction to the location change in response to κ_b. But the shift is much smaller. The shape of the price distribution is more sensitive as well, with the dispersion getting smaller (larger) in response to a larger (smaller) κ_s. The inventory distribution exhibits similar responses.

Finally, we discuss the quantitative impact of introducing idiosyncratic utility of buyers. In the

[26] Changing search-cost levels is equivalent to altering the curvature of the matching function when the functional form is urn-ball. The reason is as follows. Consider for example an urn-ball matching function with a curvature parameter $\xi > 0$ on the retail side, or $1 - e^{-\xi\theta}$. Then κ_b appears in equilibrium conditions only in the form of κ_b/xi. Increasing κ_b is therefore equivalent to proportionally decreasing ξ, and one parameter κ_b (or κ_s) is sufficient to capture search frictions in retail (or wholesale) markets.

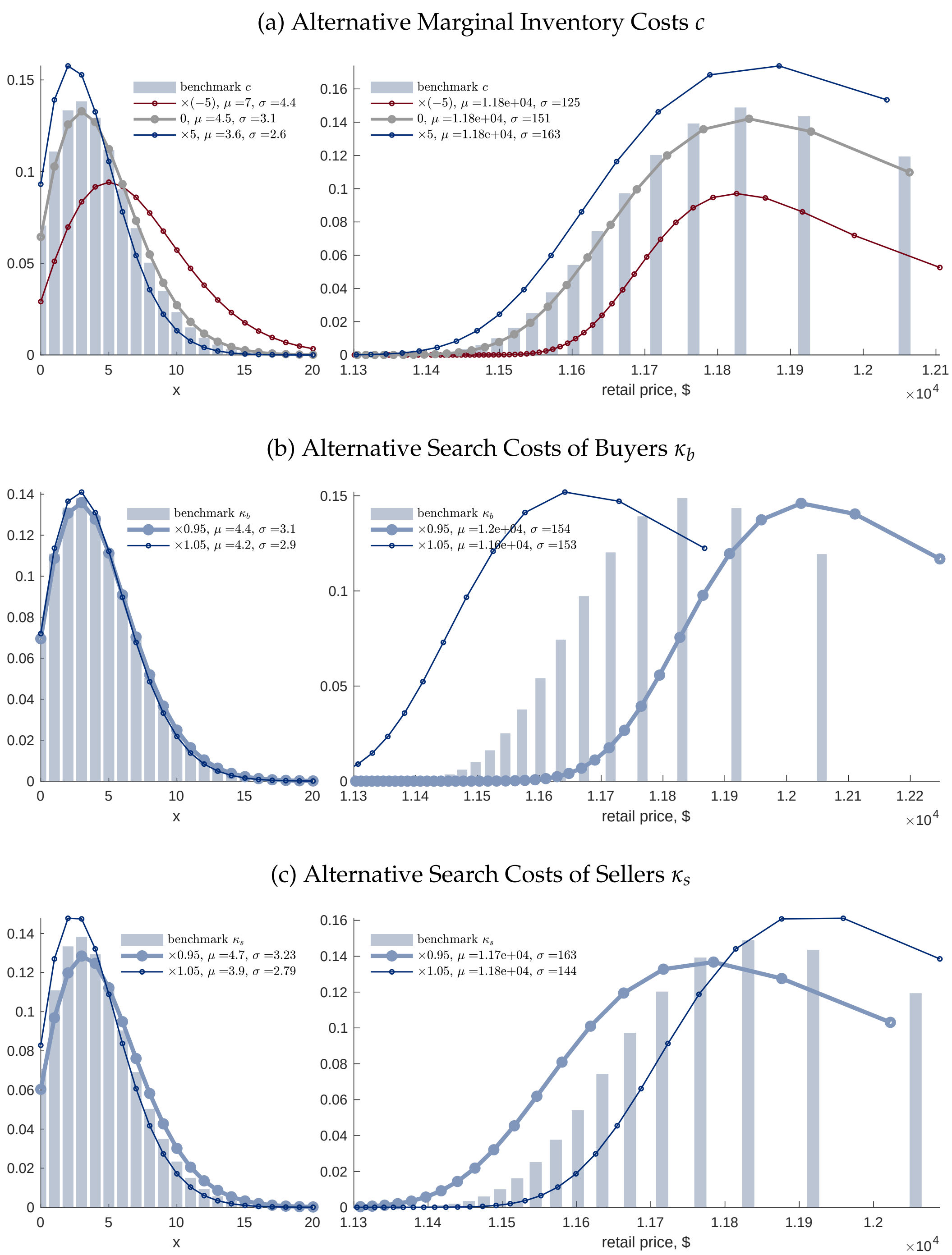


Figure C.7: Inventory and Price Distributions under Alternative Parameter Values.

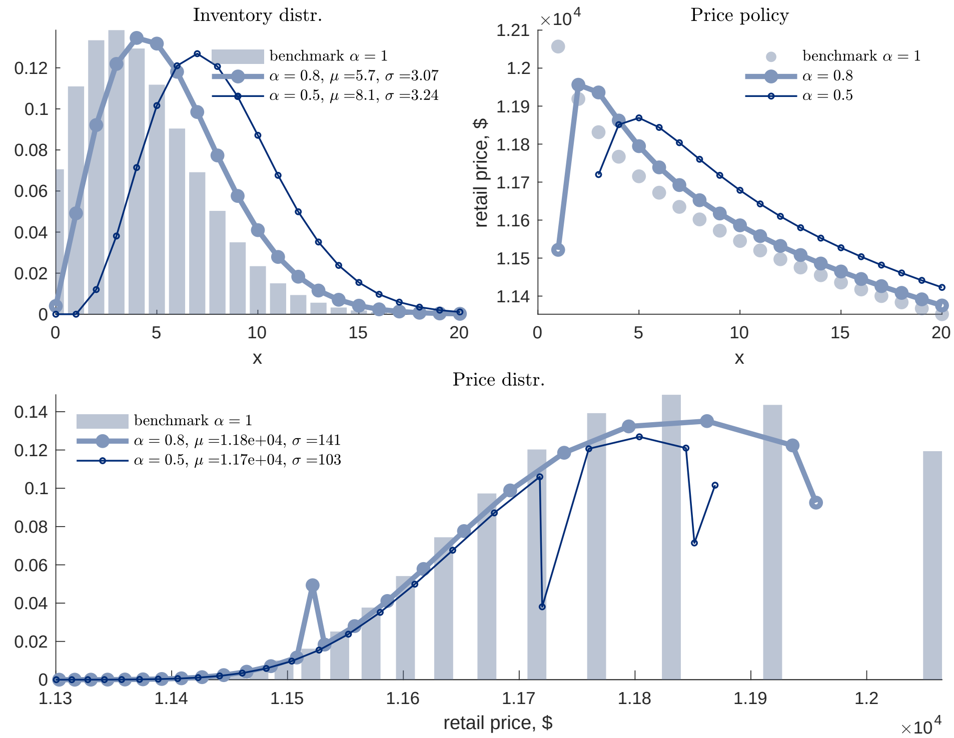


Figure C.8: Idiosyncratic Utility with Probability α of a Good Match.

data, price distribution may not always be unimodal. Indeed, introducing a nontrivial mismatch probability can jeopardize the unimodal feature of the equilibrium price distribution in our model. Figure C.8 demonstrates to what degree the existence of bad matches alters the equilibrium price distribution, and why. As discussed in section 4.2, possible bad matches incentivize dealers to increase inventory holdings, shown in the top-left panel. Price distribution exhibits additional modes when good matches are not guaranteed, plotted in the bottom panel. This is because dealers' optimal price policy is no longer monotone, shown in the top-right panel. When bad matches are more likely to occur or α becomes smaller, such incentives are stronger, but low-inventory dealers find it optimal to set low retail prices, as it is difficult to attract buyers. When inventory is sufficiently low, it is optimal for dealers not to go in any retail submarket.

References

ADAMS, C. P., L. HOSKEN, AND P. W. NEWBERRY (2011): "Vettes and lemons on eBay," *Quantitative Marketing and Economics*, 9(2), 109–127.

AGUIRREGABIRIA, V. (1999): "The dynamics of markups and inventories in retailing firms," *The review of economic studies*, 66(2), 275–308.

AKERLOF, G. A. (1970): "The market for" lemons": Quality uncertainty and the market mechanism," *The Quarterly Journal of Economics*, pp. 488–500.

ANDERSON, S. P., AND R. RENAULT (1999): "Pricing, product diversity, and search costs: A Bertrand-Chamberlin-Diamond model," *The RAND Journal of Economics*, pp. 719–735.

BAYE, M. R., J. MORGAN, AND P. SCHOLTEN (2006): "Information, search, and price dispersion," *Handbook on economics and information systems*, 1, 323–375.

BENTAL, B., AND B. EDEN (1993): "Inventories in a competitive environment," *Journal of Political Economy*, 101(5), 863–886.

BIGLAISER, G. (1993): "Middlemen as experts," *The RAND journal of Economics*, pp. 212–223.

BIGLAISER, G., AND J. W. FRIEDMAN (1994): "Middlemen as guarantors of quality," *International journal of industrial organization*, 12(4), 509–531.

BIGLAISER, G., F. LI, C. MURRY, AND Y. ZHOU (2020): "Intermediaries and product quality in used car markets," *The RAND Journal of Economics*, 51(3), 905–933.

BOUALAM, Y. (2019): "Credit Markets and Relationship Capital," Discussion paper, University of North Carolina.

BURDETT, K., AND K. L. JUDD (1983): "Equilibrium price dispersion," *Econometrica: Journal of the Econometric Society*, pp. 955–969.

BURDETT, K., S. SHI, AND R. WRIGHT (2001): "Pricing and matching with frictions," *Journal of Political Economy*, 109(5), 1060–1085.

CAILLAUD, B., AND B. JULLIEN (2003): "Chicken & egg: Competition among intermediation service providers," *RAND journal of Economics*, pp. 309–328.

CHANDRA, A., AND M. TAPPATA (2011): "Consumer search and dynamic price dispersion: an application to gasoline markets," *The RAND Journal of Economics*, 42(4), 681–704.

COEY, D., B. J. LARSEN, AND B. C. PLATT (2020): "Discounts and deadlines in consumer search," *American Economic Review*, 110(12), 3748–85.

COPELAND, A., W. DUNN, AND G. HALL (2011): "Inventories and the automobile market," *The RAND Journal of Economics*, 42(1), 121–149.

DENECKERE, R., H. P. MARVEL, AND J. PECK (1996): "Demand uncertainty, inventories, and resale price maintenance," *The Quarterly Journal of Economics*, 111(3), 885–913.

DUFFIE, D., N. GÂRLEANU, AND L. H. PEDERSEN (2005): "Over-the-Counter Markets," *Econometrica*, 73(6), 1815–1847.

GAVAZZA, A. (2016): "An empirical equilibrium model of a decentralized asset market," *Econometrica*, 84(5), 1755–1798.

GAVAZZA, A., AND A. LIZZERI (2021): "Frictions and Intermediation in Product Markets," Discussion paper, London School of Economics.

GAVAZZA, A., A. LIZZERI, AND N. ROKETSKIY (2014): "A quantitative analysis of the used-car market," *American Economic Review*, 104(11), 3668–3700.

GUO, N. (2018): "The Effect of an Early Career Recession on Schooling and Lifetime Welfare," *International Economic Review*, 59(3), 1511–1545.

HALL, G., AND J. RUST (2000): "An empirical model of inventory investment by durable commodity intermediaries," in *Carnegie-Rochester Conference Series on Public Policy*, vol. 52, pp. 171–214. Elsevier.

HARTIGAN, J. A., AND P. M. HARTIGAN (1985): "The dip test of unimodality," *The annals of Statistics*, 13(1), 70–84.

JOHRI, A., AND J. LEACH (2002): "Middlemen and the allocation of heterogeneous goods," *International Economic Review*, 43(2), 347–361.

KAPLAN, G., AND G. MENZIO (2015): "The morphology of price dispersion," *International Economic Review*, 56(4), 1165–1206.

LACH, S. (2002): "Existence and persistence of price dispersion: an empirical analysis," *Review of economics and statistics*, 84(3), 433–444.

LI, F., AND X. WENG (2017): "Efficient learning and job turnover in the labor market," *International Economic Review*, 58(3), 727–750.

MENZIO, G., AND S. SHI (2010): "Block recursive equilibria for stochastic models of search on the job," *Journal of Economic Theory*, 145(4), 1453–1494.

——— (2011): "Efficient search on the job and the business cycle," *Journal of Political Economy*, 119(3), 468–510.

MURRY, C., AND H. SCHNEIDER (2015): "The Economics of Retail Markets for New and Used Cars," in *Handbook on the Economics of Retail and Distribution*, ed. by E. Basker. Edward Elgar.

MURRY, C., AND Y. ZHOU (2020): "Consumer search and automobile dealer colocation," *Management Science*, 66(5), 1909–1934.

NAKAMURA, E., AND D. ZEROM (2010): "Accounting for incomplete pass-through," *The Review of Economic Studies*, 77(3), 1192–1230.

NOSAL, E., Y.-Y. WONG, AND R. WRIGHT (2019): "Intermediation in markets for goods and markets for assets," *Journal of Economic Theory*, 183, 876 – 906.

PETERS, M. (2000): "Limits of exact equilibria for capacity constrained sellers with costly search," *Journal of Economic Theory*, 95(2), 139–168.

PISSARIDES, C. A. (1985): "Short-run equilibrium dynamics of unemployment vacancies, and real wages," *American Economic Review*, 75(4), 676–690.

PORTEUS, E. L. (2002): *Foundations of stochastic inventory theory*. Stanford University Press.

PRESCOTT, E. C. (1975): "Efficiency of the natural rate," *Journal of Political Economy*, 83(6), 1229–1236.

REINGANUM, J. F. (1979): "A simple model of equilibrium price dispersion," *Journal of Political Economy*, 87(4), 851–858.

RHODES, A., M. WATANABE, AND J. ZHOU (2021): "Multiproduct intermediaries," *Journal of Political Economy*, 129(2), 000–000.

RUBINSTEIN, A., AND A. WOLINSKY (1987): "Middlemen," *The Quarterly Journal of Economics*, 102(3), 581–593.

RUST, J., AND G. HALL (2003): "Middlemen versus market makers: A theory of competitive exchange," *Journal of Political Economy*, 111(2), 353–403.

SALZ, T. (2017): "Intermediation and Competition in Search Markets: An Empirical Case Study," *working paper, Columbia University*.

SCARF, H. (1960): "The Optimality of (s, S) Policies in the Dynamic Inventory Problem," *Mathematical Methods in the Social Sciences*.

SCHAAL, E. (2017): "Uncertainty and unemployment," *Econometrica*, 85(6), 1675–1721.

SHEVCHENKO, A. (2004): "Middlemen," *International Economic Review*, 45(1), 1–24.

SHI, S. (2009): "Directed search for equilibrium wage–tenure contracts," *Econometrica*, 77(2), 561–584.

SMITH, E. (2004): "Intermediated search," *Economica*, 71(284), 619–636.

SPULBER, D. F. (1996): "Market microstructure and intermediation," *Journal of Economic perspectives*, 10(3), 135–152.

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

STIGLER, G. J. (1961): "The economics of information," *Journal of political economy*, 69(3), 213–225.

WHITIN, T. M. (1955): "Inventory control and price theory," *Management science*, 2(1), 61–68.

WOLINSKY, A. (1986): "True monopolistic competition as a result of imperfect information," *The Quarterly Journal of Economics*, 101(3), 493–511.

WRIGHT, R., P. KIRCHER, B. JULIEN, AND V. GUERRIERI (2017): "Directed search: A guided tour," Discussion paper, National Bureau of Economic Research.

WRIGHT, R., AND Y.-Y. WONG (2014): "Buyers, Sellers, And Middlemen: Variations On Search-Theoretic Themes," *International Economic Review*, 55(2), 375–397.

ZETTELMEYER, F., F. S. MORTON, AND J. SILVA-RISSO (2006): "Scarcity rents in car retailing: Evidence from inventory fluctuations at dealerships," Discussion paper, National Bureau of Economic Research.