The Price Theory and Empirics of Inventory Management*

Fei Li Charles Murry Can Tian Yiyi Zhou†

November 10, 2019

Abstract

We develop a directed search model where buyers purchase goods produced by sellers through intermediaries. The presence of search frictions creates demand uncertainty and makes instantaneous replenishment impossible. To avoid the risk of stockout, an intermediary holds inventory. The intermediary's trade-off between inventory cost and stockout risk depends on the size of inventory and determines its optimal retail pricing and restocking policy. In equilibrium, when the intermediary's inventory increases, he posts a lower retail price to speed up sales and depresses wholesale price to slow down purchases. In the steady state, the equilibrium generates unimodal distributions of both inventory holdings and retail prices. Using a dataset that contains detailed information on used car listings, we empirically examine dealers' inventory, new orders, sales, and retail prices. Our empirical findings are consistent with the model predictions.

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

JEL Classification Codes: D82, D83, L15, L62

*We thank Gary Biglaiser, Michael Choi, Luca Flabbi, Chao Fu, Qing Gong, Veronica Guerrieri, Adrian Masters, Brian McManus, Peter Norman, Stan Rabinovich, Andrei Shevchenko, Eric Smith, Makoto Watanabe, Jonathan Williams, Randall Wright, Valentin Verdier, Jidong Zhou, Qiankun Zhou and participants of 2019 Midwest Macroeconomic conference for helpful comments.

†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.

Contents

1 Introduction

Dealers, retailers, and other intermediaries play a prominent role in well-functioning markets. Their rationale stems from various market frictions suggested in the literature. Because of these frictions, intermediaries inevitably face uncertain demand, which in turn leads to inventory challenges. Managing inventory while facing uncertain demand requires intermediaries to (1) manage new orders and make scrappage decisions and (2) manage revenue and sales through dynamic pricing policies. In this paper, we examine how inventory control and revenue management shape the equilibrium price dynamics and cross sectional price dispersion in an environment where intermediaries face competition.

We show that revenue and inventory management in the face of uncertain demand generates patterns of price dispersion consistent with real markets in the absences of ex-ante heterogeneity of agents. In particular, we predict unimodal price distributions and patterns of inventory management consistent with data on used car dealers, and consistent with other recent empirical studies of price dispersion.[1] We borrow elements of models from the on-the-job search literature (Shi, 2009; Menzio and Shi, 2010, 2011) to model consumer search and intermediaries' inventory management. 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 retail buyers observe all retail prices and decide the set of intermediaries to seek, and in wholesale markets, sellers observe all requested orders and decide the set of intermediaries to search for. Such a 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 equilibrium retail pricing rules and stocking decisions of the intermediary, which depend on the current inventory size. The equilibrium implies an ideal inventory size. If the intermediary's inventory size is large, the risk of stockout is small, so it can afford to offer a low wholesale price and to make new orders slowly, but it has the incentive to charge a low retail price to boost the speed of retail sales. On the contrary, if the intermediary's inventory size is small, the risk of stockout is large, so it has the urgent to make new orders quickly at a high wholesale price, but it also finds it optimal to make retail sales only at a high retail price. Overall,

[1] See Lach (2002), Hosken and Reiffen (2004), Reuter and Caulkins (2004), Woodward and Hall (2012), Kaplan and Menzio (2015), Kaplan, Menzio, Rudanko, and Trachter (2019) as a few examples.

the inventory size of an intermediary determines its optimal policy, and therefore, the law of motion of its inventory. In *steady state*, a histogram of inventory levels must be mostly concentrated around the ideal level and gradually decreases when moving away from the ideal level. Because the retail price is a monotone function of the inventory, the steady state distribution of retail price must also be single-peaked in the equilibrium. Furthermore, because inventory is costly to seek, the optimal policy follows an (s, S) rule – for example, if the marginal benefit of restocking is sufficiently low, then the intermediary will allow small deviations from the ideal inventory. We extend the model to allow for product differentiation and multiple wholesale units.

We test whether our theory of inventory management and dynamic pricing is consistent with the used car industry. The used car industry is an ideal laboratory to study these issues because car dealers face large inventory costs, inventory decisions do not involve long-term contracts, dealers are able to quickly adjust prices, the wholesale market is relatively liquid and dealers can typically make weekly stocking decisions, and a majority of used cars sell though dealers. First, we present a descriptive analysis of the dynamic behavior of retail price, inventory, and new orders. Second, we conduct statistical tests of the predictions of our model in the context of reduced-form estimations of the optimal decision rules of intermediaries, examining how inventory affects retail pricing, the quantity of sales, and the quantity of new orders. In our analysis, we control for both observed and unobserved fixed effects of the car and the dealer, and other potential sources of spurious state dependence such as seasonality. We find the data is consistent with our theoretical implications. Third, we test whether the distribution of the deviation from the ideal level inventory of intermediaries are single peaked, and whether the residual price dispersion after controlling observable characteristics are single peaked as predicted in our theoretical model. We find broad support of the model prediction from our empirical evidence.

Related Literature and Contribution. Our paper contributes to the literature of search and price dispersion. The idea that the law of one price fails due to imperfect information was first suggested by Stigler (1961) and formalized in equilibrium models in Varian (1980), Burdett and Judd (1983) and Stahl (1989). These models generate price dispersion in mixed-strategy equilibria where firms randomly choose prices. These papers inspired a cast literature, including a line of research on consumer search models. These models typically feature oligopolistic price competition and (essentially) require consumers with heterogeneous information, or firms with heterogeneous cost or visibility to generate price dispersion.[2] See Baye, Morgan, and Scholten (2006) for a survey. Although in all aforementioned consumer search papers, intermediaries are abstracted

[2]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) and Chandra and Tappata (2011).

away, Janssen and Shelegia (2015) and Garcia, Honda, and Janssen (2017) introduce retailers into consumer search frameworks and show that price dispersion can be generated without any ex ante heterogeneity of firms or consumers. However, all the aforementioned models generate non-unimodal price dispersion, despite it has been well documented that price dispersion in many markets is unimodal. Our model generates a *unimodal* price dispersion due to inventory dynamics without assuming any ex ante heterogeneity among buyers or sellers. Therefore, our paper complements this literature by highlighting the importance of inventory management as a realistic channel to accommodate the shape of price dispersion in the data. Moreover, since price dispersion is obtained in a mixed-strategy equilibrium in aforementioned models, a typical test is to examine if there is within-distribution price rank reversals among firms (Chandra and Tappata 2011). Our theory provides an alternative justification of within-distribution price rank reversals due to (unobservable) dynamics of inventory.

Our model of consumer search and inventory management is more reminiscent of the rich labor literature on job search and wage dispersion, and our inventory problem generalizes the on-the-job search worker allocation problem. See Burdett and Mortensen (1998) and Fu (2011) as examples and Rogerson, Shimer, and Wright (2005) for a survey. In contrast to typical consumer search models, labor search models typically focus on search dynamics (crucial for us thinking about dynamic inventory management) and generate ex-post heterogeneity that generates wage dispersion (or, in our case, price dispersion). We model search as directed in the sense of Moen (1997) and Burdett, Shi, and Wright (2001). We refer readers to Wright, Kircher, Julien, and Guerrieri (2017) for a comprehensive survey of the literature. Our model is most closely related to the framework developed by Shi (2009), Gonzalez and Shi (2010), and Menzio and Shi (2010, 2011). Watanabe (2010, 2018) study intermediaries in directed search framework, and focus on the endogenous emergence of intermediaries. Watanabe (2019) uses a similar framework to explain the empirical relation between intermediaries' price premium and their advantage in inventory holding. On the contrary, we focus on the impact of the inventory dynamics on pricing.

We are able to generate a unimodal price distribution by introducing an inventory management problem of an intermediary, or retailer. There is a long tradition of attributing the existence of intermediaries to search frictions (Rubinstein and Wolinsky, 1987).[3] While most papers above focus on search frictions per se and do not allow intermediaries to hold multiple inventories, 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. Rhodes, Watanabe, and Zhou (2018) introduce multiple products and study the optimal portfo-

[3] Alternatively, another literature emphasizes asymmetric information as the primary source of market frictions and argue that intermediaries mainly serve as information providers or certifiers. See Biglaiser (1993), Lizzeri (1999), Biglaiser and Li (2018) as a few examples.

lio choice of intermediaries. There are many recent empirical papers about intermediaries and search frictions (Gavazza, 2016; Hendel, Nevo, and Ortalo-Magné, 2009; Salz, 2017), but we are not aware of any empirical study on the role of intermediaries inventory management problem and price dispersion.

To be sure, there is a large literature in operations research that study the optimal inventory and revenue management decision problem. See Talluri and Van Ryzin (2006) and Porteus (2002) for textbook treatment. The idea to combine pricing theory and inventory control was first proposed by Whitin (1955), but 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 Bental and Eden (1993) who use a demand-uncertainty model a la Eden (1990) and obtain a similar relation between inventories and retail prices as our model. On the empirical side, 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. Aguirregabiria (1999) highlights the role of inventory on pricing, and he argues that the stock-out probability and the fixed cost of ordering can explain supermarket discounts/sales behavior.

Organization. The rest of the paper is organized as follows. Section 2 introduces the theoretical model, characterize the equilibrium, and derive empirical implications. Section 4 describes the dataset and present descriptive analysis of the dynamic behavior of inventories and retail prices, and reduced form regression of the intermediaries' optimal decision on pricing, and new orders and its consequences. Also, it test whether the distributions of normalized inventory and residual price are unimodal.

2 Model

In this section, we develop a stylized search and matching model where buyers purchase goods produced by sellers through intermediaries. We assume a search friction that creates demand uncertainty and makes instantaneous replenishment of inventory impossible for the intermediary. To avoid the risk of a stockout, an intermediary holds inventory. The intermediary's trade-off between the risk of stockout and the cost of purchasing and holding inventory depends on the size of inventory and determines his optimal retail pricing and purchase policy. We show that the model admits a unique equilibrium and a unimodal steady-state distribution of retail price.

2.1 Environment

Time is discrete and lasts forever, $t = 0, 1, 2, \ldots$. The economy is populated by buyers, sellers and intermediaries.

Buyers and Sellers. In each period, 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 cannot meet directly. The consumption goods can only be delivered from sellers to buyers through a third party.

Intermediaries. There is a unit measure of ex ante identical long-lived intermediaries, 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 only sell if his current inventory is positive. To avoid the risk of stockout, an intermediary can hold inventory for a cost. Specifically, holding x units of inventory costs an intermediary $c(x)$ per period for $x = 0, 1, 2, \ldots$. The cost function $c : \mathbb{N} \to \mathbb{R}$ is increasing and convex, such that $c(0) = 0$.[4] All intermediaries are risk neutral and share a common discount factor $\delta \in (0, 1)$.

Let $g_t : \mathbb{N} \to [0, 1]$ be the probability mass function of the distribution of inventory holding across intermediaries in period 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 submarket indexed by the wholesale price $w \in \mathbb{R}$. In each wholesale submarket w, the ratio of sellers to intermediaries is denoted by $\lambda(w)$. We refer to $\theta(p)$ and $\lambda(w)$ as the *tightness* of the corresponding retail and wholesale submarkets.

Search and Matching. Search is directed. In each period, an intermediary can choose to enter at most one submarket without incurring any explicit search cost. In this way, we capture the intermediary's retail/wholesale pricing problem as a choice of the corresponding submarkets. With probability 1/2, he can choose a retail submarket to enter and search for buyers; with probabil-

[4] We adopt the convention that $0 \in \mathbb{N}$ henceforth. Here, convexity of a discrete function $c(\cdot)$ means that $c(x+1) - c(x)$ increases in $x, \forall\, x \in \mathbb{N}$.

ity 1/2, he can choose a wholesale submarket to enter and search for sellers.[5] A buyer sees all the retail submarkets (prices) and chooses to enter at most one retail submarket in each period by paying a search cost $\kappa_b > 0$; while a seller sees all wholesale submarkets (prices) chooses to enter at most one wholesale submarket in each period by paying an entry cost $\kappa_s > 0$. When an intermediary and a buyer meet in retail submarket p, the buyer buys one unit of 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.

Submarkets are frictional. In each retail submarket and in each period, a standard matching function, $\mathcal{M}_r(b, m)$, determines the measure of bilateral meetings as a function of the measures of buyers and intermediaries in this submarket, b and m. We assume that the matching function is strictly increasing, strictly concave, homogeneous of degree one, and continuously differentiable. The total number of meetings cannot exceed the number of traders on the short side of the market, so $\mathcal{M}_r(b, m) \leq \min\{b, m\}$. Also, we assume that $\mathcal{M}_r(0, m) = \mathcal{M}_r(b, 0) = 0$ for any $b, m \in \mathbb{R}$, and $\lim_{\theta\to\infty} \mathcal{M}_r(\theta, 1) = 1$, $\lim_{\theta\to\infty} \mathcal{M}_r(\theta, 1)/\theta = 0$ and $\lim_{\theta\to 0} \mathcal{M}_r(\theta, 1)/\theta = 1$.

Since the matching function is constant return to scale, in every submarket, an agent's matching probability can be expressed as a function of the tightness of the submarket. Specifically, consider a retail submarket p. In each period, an intermediary meets a buyer with probability

$$\mathcal{M}_r(b, m)/m = \mathcal{M}_r(\theta(p), 1) = \phi_r(\theta(p)), \theta(p) \equiv b/m,$$

where $\phi_r : \mathbb{R}_+ \to [0, 1]$ is a bounded, twice-differentiable, strictly increasing and strictly concave function such that $\phi_r(0) = 0$ and $\lim_{\theta\to\infty} \phi_r(\theta) = 1$. On the other side, a buyer meets an intermediary with probability

$$\mathcal{M}_r(b, m)/b = \mathcal{M}_r(1, 1/\theta(p))) \equiv \psi_r(\theta(p))$$

where $\psi_r : \mathbb{R}_+ \to [0, 1]$ is a bounded, twice-differentiable, strictly decreasing function such that $\psi_r(0) = 1$, $\lim_{\theta\to\infty} \psi_r(\theta) = 0$, and $\psi_r(\theta) = \phi_r(\theta)/\theta, \forall\theta > 0$.

Similarly, in each wholesale submarket and in each period, the measure of bilateral meetings is also determined by a matching function $\mathcal{M}_w(s, m)$ where s and m denote the measure of sellers and intermediaries in the submarket, and an agent's matching probability can also be expressed as a function of the tightness of the submarket. Therefore, in a wholesale submarket w, an intermediary meets a sellers with probability $\phi_w(\lambda(w))$ where $\phi_w(\cdot) \equiv \mathcal{M}_w(\cdot, 1)$ is a bounded, twice-differentiable, strictly increasing and strictly concave function such that $\phi_w(0) = 0$ and $\lim_{\lambda\to\infty} \phi_w(\lambda) = 1$. On the other side, a seller meets an intermediary with probability $\psi_w(\lambda(w))$

[5] This specification is made to ease the presentation of the intermediary's dynamic programming. Our analysis does not crucially depend on it.

where $\psi_w : \mathbb{R}_+ \to [0,1]$ is a bounded, twice-differentiable, strictly decreasing function such that $\psi_w(0) = 1$, $\lim_{\lambda\to\infty} \psi_r(\lambda) = 0$, 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 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. Second, 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, the measure of active buyers and sellers are pinned down by free-entry conditions. These assumptions allow us to focus on the dynamic problem of intermediaries and to maintain tractability. 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. Third, we assume that a seller supplies at most one unit. In some applications, this assumption seems restrictive. The assumption is made for simplification. In Section 3, we discuss an extension where sellers can supply a multi-unit package. Lastly, we assume that an intermediary can visit either a retail submarket and a wholesale submarket with equal probability. This assumption is made to simplify the exposition. Alternatively, one can allow an intermediary to enter exactly one submarket in each period, which can be either a retail submarket or a wholesale submarket. One can also divide every period into two sub-period, day and night. Wholesale submarkets are operated in the day; while retails submarkets are operated in the night. As we will argue in Section 2.5, our main empirical implications do not depend on the choice of these specifications.

2.2 Individual Problems and Definition of Equilibrium

The Intermediary's Problem. Consider an intermediary with inventory x. His lifetime expected profit is given by

$$\begin{aligned} V(x,g) &= \max_{p,w} -c(x) + \tfrac{1}{2}\phi_r(\theta(p))[p + \delta V(x-1,g')] + \tfrac{1}{2}\phi_w(\lambda(w))[-w + \delta V(x+1,g')] \\ &+ \left[1 - \tfrac{1}{2}\phi_r(\theta(p)) - \tfrac{1}{2}\phi_w(\lambda(w))\right]\delta V(x,g') \end{aligned} \tag{1}$$

for every $x \geq 0$. The intermediary's value function has two state variables. An individual state, which is his current inventory size x, and an aggregate state, which is the distribution of inventory holding across intermediaries g. In the current period, the intermediary incurs inventory cost

$c(x)$ regardless of his choice. Given each pair (p, w), the intermediary enters retail submarket p and wholesale submarket w with equal probability. If the intermediary searches for buyers, he meets a buyer with probability $\phi_r(\theta(p))$, sells one unit of the good at price p, and his next period inventory size becomes $x - 1$. If the intermediary searches for sellers, he meets a seller with probability $\phi_w(\lambda(w))$, buys one unit of the good at price w, and his next period inventory size becomes $x + 1$. If the intermediary meets neither a buyer nor a seller, his inventory size remain x but the aggregate state changes to g'. As it will become clear soon, the aggregate state is irrelevant for the intermediary's decision, so we delegate its law of motion to section 2.4. We impose a boundary condition $V(-1, g) = \underline{V}$, which is sufficiently small for any g to capture the idea that an intermediary cannot sell when stockout ($x = 0$).

The Problem of Buyers and Sellers. In each period, a buyer decides whether and where to 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 probability that the consumer meets an intermediary $\psi_r(\theta(p))$ and the the surplus from buying the good at price p. Moreover, if the consumer-to-intermediary ratio 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 must be 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 probability that 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 follow-

ing conditions: (1) $V(x)$ and $p(x), w(x)$ solve (1) for every g; and (2) the retail market tightness function $\theta(p)$ satisfies (2), and wholesale market tightness function $\lambda(w)$ satisfies (3).

Our solution concept, developed by Shi (2009) and Menzio and Shi (2010, 2011), is similar to the standard recursive competitive equilibrium where strategies of each agent are optimal given the strategies of other agents except that the agent's problem does not depend on the distribution of heterogenous individual states across agents g. In general, the aggregate state g varies over time and it is natural to believe 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.

2.3 Equilibrium Characterization

Because search is directed, an intermediary faces a trade-off between the probability of trade 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 featured by 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 $\theta(p)$ must satisfy

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

where the second equality holds due to the fact that $\phi_r(\theta) = \psi_r(\theta)/\theta, \forall \theta > 0$. 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

$$V(x) = -c(x) \quad + \quad \underbrace{\frac{1}{2}\left[\max_{\theta \geq 0} \phi_r(\theta)[u + \delta V(x-1) - \delta V(x)] - \kappa_b \theta\right]}_{\text{retail transaction gain}}$$
$$+ \quad \underbrace{\frac{1}{2}\left[\max_{\lambda \geq 0} \phi_w(\lambda)\delta[V(x+1) - V(x)] - \kappa_s \lambda\right]}_{\text{wholesale transaction gain}} + \delta V(x). \tag{6}$$

That is, the intermediary's retail and whole 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 the decision problem (6). The above argument also makes it clear why the aggregate state is redundant to solve the equilibrium. An intermediary wants to maximize the expected life-time total utility that he delivers to consumers, 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 aggregate state 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, the trade-off between the price and the transaction speed disappears. The discussion above is summarized by the following proposition. The formal argument is essentially identical to the one in Menzio and Shi (2010, 2011), so it is omitted.

Proposition 1. *An equilibrium exists and it is the unique recursive equilibrium.*

Now we are ready to characterize the unique equilibrium. From (6), it follows that the optimal

$\theta^*(x)$ must satisfy the following first-order condition (FOC):

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

and $\theta^*(x) \geq 0$ with complementary slackness. The FOC in (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

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

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, $u + \delta[V(x-1) - V(x)]$ and $\delta[V(x+1) - V(x)]$, 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. The proof is relegated to Appendix A, and we provide the intuition here. The result is driven by two assumptions. First, 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. Because the risk of stockout is strictly decreasing in the intermediary's inventory size x, the marginal benefit of holding inventory is decreasing in x. Second, the cost of inventory holding is convex, which further diminishes the marginal benefit of inventory holding. 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 $\lambda^*(x)$ decreases in x, and*

2. *both the retail price $p^*(x)$ and the wholesale price $w^*(x)$ decrease in x.*

Proposition 2 is one of the main empirical implications of the model. The intuition behind is simple. There is some "ideal size" of stock, whenever the actually stock deviates from the mean size, the intermediary adjusts the sales and ordering strategy to make the inventory regress toward the mean size. Specifically, when an intermediary's stock is too high, he lowers both the retail price and the wholesale price. In doing so, he is more likely to sell goods and less likely to buy goods, reducing his next period inventory holding in expectation. When his stock is too low, the intermediary raises both the retail price and the wholesale price to increase his next period inventory holding in expectation.

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*. 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 (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 $x > s$.

2.4 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:

$$\begin{aligned} g_{t+1}(x) &= g_t(x-1)\frac{\phi_w(\lambda^*(x-1))}{2} + g_t(x+1)\frac{\phi_r(\theta^*(x+1))}{2} \\ &+ g_t(x)[1 - \frac{\phi_r(\theta^*(x))}{2} - \frac{\phi_w(\lambda^*(x))}{2}] \end{aligned} \tag{10}$$

for every $x = 0, ..., s$. The left-hand side of (10) is the measure of intermediaries who hold x units of inventory in period $t+1$. The right-hand side of (10) has three parts. First, $g_t(x-1)/2$ of intermediaries hold $x-1$ units of inventory and search in a wholesale submarket with tightness $\lambda^*(x-1)$ in period t, and $\phi_w(\lambda^*(x-1))$ of them find sellers, trade, and increase their period $t+1$ stock to x. Second, $g_t(x+1)/2$ of intermediaries hold $x+1$ units of inventory and search in a retail submarket with tightness $\theta^*(x+1)$ in period t, and $\phi_r(\theta^*(x+1))$ of them find buyers, trade, and decrease their period $t+1$ stock to x. Third, $g_t(x)$ of intermediaries hold x units of inventory in period t and $1 - \phi_r(\theta^*(x))/2 - \phi_w(\lambda^*(x))/2$ of them fail to meet either buyers or sellers, so their period $t+1$ inventory remain x.

At steady state $g_t = g_{t+1}$, 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.[6] The intuition behind is very simple. By Proposition 2, whenever an intermediary's inventory deviates from the "ideal level" denoted by x^*, he adjusts the retail or wholesale policy θ and λ to push the future stock back to the x^*. The more the stock deviates from the mean level, the faster the speed of the regression is. Therefore, in the state steady, the mass at the ideal level of inventory x^* is the

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

highest, and the mass monotonically decreases as the stock becomes farther and farther away from the mean level.

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 the ideal level x^*, one should expects the intermediary's retail price is equal to or close to $p^*(x^*)$ for most 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 of search models that relies 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* (relative positions of intermediaries varying over time within the price distribution) which is also consistent with a number of empirical studies such as Lach (2002) and Chandra and Tappata (2011). In the literature, such a phenomena 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.

2.5 On Specification of Market Choice

Before moving forward, we want to clarify the simplification on the intermediary choice of submarkets. We assume that an intermediary can enter retail submarkets and wholesale submarkets with equal probabilities. The assumption is not essential. Alternatively, one can assume that an intermediary can still enter exactly one submarket in each period, but the probability of entering a retail submarket and a wholesale submarket is endogenous, denoted by $\eta(x) \in [0,1]$. The intermediary's problem is therefore divided into two steps. First, given its inventory x, the intermediary chooses to enter some retail submarket with probability $\eta(x)$ and to enter some wholesale submarket with the complementary probability. Given the realization of the first step randomization, the intermediary chooses which retail or wholesale submarket to enter. It is easy to see that

the optimal $\eta(x)$ is a bang-bang solution, and the solution depends on the difference between the retail transaction gain and wholesale transaction gain in equation (6). When x is sufficiently high, the gain from retail transaction dominates, so $\eta(x) = 1$; when x is sufficiently low, the gain from wholesale transaction dominates, and $\eta(x) = 0$. By Lemma 1, the retail transaction gain is increasing in x the wholesales transaction gain is decreasing in x; and thus there is a cutoff x^* such that $\eta(x) = 1$ if $x \geq x^*$ and $\eta(x) = 0$ otherwise. It is easy to see that the value function $V(x)$ and the policy functions $\theta^*(x)$ and $\lambda^*(x)$ of the intermediary's problem, the law of motion of equilibrium inventory dynamics, and the unimodality of the steady state distribution of inventory do not qualitatively change. Therefore, we consider the endogenous probability choice $\eta(\cdot)$ redundant, and keep it exogenous.

3 Applications

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.

3.1 Heterogenous Intermediaries

Intermediaries may be heterogenous 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 manufactures, 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 heterogeneity can lead to different size, sales and profitability among intermediaries.

One can extend our model by allowing heterogenous 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 problem

$$\begin{aligned} V_j(x) = -c_j(x) \quad &+ \quad \tfrac{1}{2}\left[\max_{\theta \geq 0} \phi_{r,j}(\theta)[u + \delta V_j(x-1) - \delta V_j(x)] - \kappa_b \theta\right] \\ &+ \quad \tfrac{1}{2}\left[\max_{\lambda \geq 0} \phi_{w,j}(\lambda)\delta[V_j(x+1) - V_j(x)] - \kappa_s \lambda\right] + \delta V_j(x). \end{aligned} \tag{12}$$

The rest of the analysis is identical to the one in the baseline model. 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.

3.2 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

$$\begin{aligned} V(x) = -c(x) \quad &+ \quad \frac{1}{2}\left[\max_{\theta \geq 0} \phi_r(\theta)[u + \delta V(x-1) - \delta V(x)] - \kappa_b \theta\right] \\ &+ \quad \frac{1}{2}\left[\max_{\lambda \geq 0, y \in \mathbb{N}} \phi_w(\lambda)\delta[V(x+y) - V(x)] - \kappa_s \lambda\right] + \delta V(x). \end{aligned} \tag{13}$$

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

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

The rest of the equilibrium analysis is straightforward. What's interesting is that the model can generate the classic (s, S)-rule in 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 order and w_1 is the linear unit price. In our model, we do not impose any structure on the wholesale pricing rule. Instead, 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 $x \leq s$ is sufficiently small, it is optimal to set $\lambda^*(x) > 0$. Also, because the marginal product cost is zero, it is immediate that $y^*(x) = S - x$.

3.3 Horizontal Product Differentiation

In reality, buyers have idiosyncratic tastes. To capture this idiosyncratic utility, we introduce horizontal product differentiation a la Perloff and Salop (1985) and Wolinsky (1986) into the model: 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

$$\kappa_b \geq \psi_r(\theta(p)) \int \max\{\tilde{u} - p, 0\} dG(\tilde{u})^x \tag{14}$$

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.[7] 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

$$\begin{aligned} V(x) = -c(x) \quad &+ \quad \tfrac{1}{2}\left[\max_{\theta \geq 0} \phi_r(\theta)\Phi(x)[u + \delta V(x-1) - \delta V(x)] - \kappa_b \theta\right] \\ &+ \quad \tfrac{1}{2}\left[\max_{\lambda \geq 0} \phi_w(\lambda)\delta[V(x+1) - V(x)] - \kappa_s \lambda\right] + \delta V(x). \end{aligned} \tag{15}$$

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

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

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

[7] Cachon, Gallino, and Olivares (2018) document that expanding variety across models significantly increase car dealers' sales.

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

$$\kappa_b \geq \psi_r(\theta(p))\Phi(x)(u-p)$$

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.

3.4 Vertical Product Differentiation

High and low quality goods, and buyers' reservation value

$$u_H > u_L > 0$$

A dealer holds x_H units of H-good and x_L units of L-good

$$
\begin{aligned}
V(x_H, x_L) &= \tfrac{1}{4}\left[\max_{\theta_H \geq 0} \phi_r(\theta_H)[u_H + \delta V(x_H - 1, x_L) - \delta V(x_H, x_L)] - \kappa_b \theta_H\right] \\
&= \tfrac{1}{4}\left[\max_{\theta_L \geq 0} \phi_r(\theta_L)[u_L + \delta V(x_H, x_L - 1) - \delta V(x_H, x_L)] - \kappa_b \theta_L\right] \\
&+ \tfrac{1}{4}\left[\max_{\lambda_H \geq 0} \phi_w(\lambda_H)\delta[V(x_H + 1, x_L) - V(x_H, x_L)] - \kappa_s \lambda_H\right] \\
&+ \tfrac{1}{4}\left[\max_{\lambda_L \geq 0} \phi_w(\lambda_L)\delta[V(x_H, x_L + 1) - V(x_H, x_L)] - \kappa_s \lambda_L\right] \\
&+ \delta V(x_H, x_L) - c(x_H + x_L)
\end{aligned}
$$

Proposition 4. *In the equilibrium,*

1. *Both $V(x_H, x_L) - V(x_H - 1, x_L)$ and $V(x_H, x_L) - V(x_H, x_L - 1)$ decreases in x_H and x_L.*
2. *Both $\theta_H(x_H, x_L)$ and $\theta_L(x_H, x_L)$ increase in x_H and x_L.*
3. *Both $\lambda_H(x_H, x_L)$ and $\lambda_L(x_H, x_L)$ decrease in x_H and x_L.*

4 Application to Used-Car Market: Empirics

In this section we examine data on a large number of used car listings. The data allow us to construct inventories for dealers, and we observe listing prices, stocking decisions, and sales. We examine how the inventory affects dealers' decisions of placing new orders, sales, and retail price setting. Our findings are consistent with the predictions of the model.

4.1 Used Car Market

A number of factors make the used car market suitable for our study. First, the market is highly decentralized, and used cars are heterogenous, making the search and matching frictions in the market non-trivial. As a result, many transactions are intermediated. Nationally, about two-thirds of used car sales are made by dealers. Second, inventory costs of used car dealers is significant. Dealers must manage both value erosion as assets age and holding costs which include floor-plan inventory investment and cost of capital.[8] Third, cars are durable goods. Most buyers and sellers do not make frequent purchase, so it is uncommon to sign long-term contract with

[8] 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.

dealers. Fourth, stocking decisions can be made frequently. Dealers acquire cars form individuals, as well as in wholsale auctions and from other dealers. Fifth, dealers frquently adjust prices. These features suggest that the interaction between inventory control and search friction is important in the used-car market, making the prediction of our theory applicable.[9]

4.2 Data

We obtain information on daily used car listings from a large car listings platform, cars.com. We observe the daily listings for each dealer who used the platform in the state of Ohio in 2017 (52 weeks in total). 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, trim with a particular set of options, exterior color, odometer mileage, whether it is certified by the dealer, the daily listing price from the date when it was initially listed to the date when it was removed from the website.

From the listing data, we count the numbers of new listings, listing removals (including sales), and inventory in each week. Figure 1 displays the three statistics over time. Overall inventories fluctuate between 42 and 46 thousand units in our sample of dealers (the second panel) and there is substantial aggregate variation in the activity of new listings and removals.

Products and dealers. While our model analyzes homogenous products, in reality, cars are highly differentiated. To better match the data and the model, we group cars into different car types based on their body style and age. We treat each car type as a product. We only include non-luxury sedans and SUVs younger than 20 years old.[10] 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 leave us with eight product categories in total.

We drop the dealers who listed fewer than 10 cars throughout the whole year. Those dealers are considered to be too small to be representative. This criterion deletes 263 product-dealer observations. In the end, our sample includes 713 dealers. Figure B.1a in Appendix B presents the number of cars listed by each dealer in an average week. 16% of all dealers listed 10-20 cars, and 46% hold fewer than 40 cars in an average week. A dealer typically sells multiple products. Among all dealers in our sample, 80% of them sell all eight products that we focus on. Since our model assumes that every intermediary sells a single product, we treat a dealer selling n types of

[9] 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

[10] Non-Luxury brands include Chevrolet, Chrysler, Dodge, Ford, GMC, Honda, Hyundai, Jeep, Kia, Mazda, Mercury, Mitsubishi, Nissan, Pontiac, Saturn, Subaru, Suzuki, Toyota, and Volkswagen.

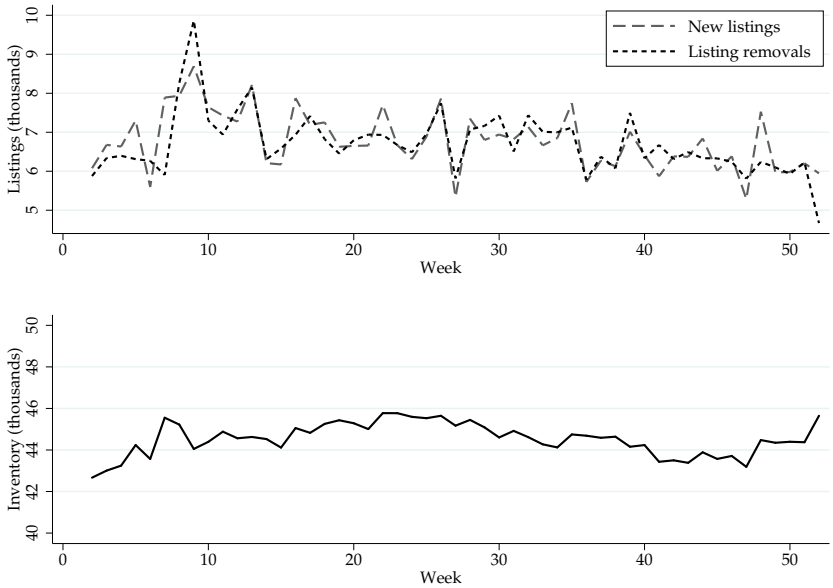


Figure 1: Inventory, New Listings, and Listing Removals (Thousands)

Note: The dash, short dash, and solid lines depict the number of total new listings, the number of listing 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 1: Descriptive Statistics: Car Level

	Mean	SD	Q10	Median	Q90
Listing weeks	6.67	6.48	1	5	14
Car age	5.58	3.68	2	4	11
Initial listing price ($)	16,007	7,990	6,999	14,972	25,990
Last listing price ($)	15,345	7,744	6,935	14,000	24,990
Price change ($)	-663	1,371	-2,004	-50	0

Note: Sample selection is described in text. The sample includes 332,459 listings. Data source: Cars.com.

products as n distinct dealers. In the rest of the paper, we consider each product-dealer combination as a distinct economic agent.

Sample for analysis. Some dealers sell some products only for short periods (even just one week). For those product-dealer combinations, the infrequency of positive orders is not associated with an inventory decision, but it is the result of the dealer's decision that the product will not be sold in the future. If we considered these product-dealer combinations in our working sample we would introduce a spurious upward bias in the frequency of zero orders and thus an downward bias in our estimate of the impact of inventory on orders. In order to avoid this problem, we concentrate our analysis on the set of product-dealers that have positive inventory at every week in 2017. As a result, we drop 3,689 product-dealers and end up with a balanced panel of 4,416 product-dealers. Due to this criterion, we drop roughly 30% of product-dealer-week observations and end up with 229,632 product-dealer-week observations representing 332,459 cars that were listed by 713 dealers. In total, it has 2,141,236 car-week observations. The top five brands take around 65% market share in total, and their market shares are Chevrolet (20.49%), Ford (17.57%), Toyota (9.31%), Honda (9.08%), and Jeep (8.65%). Table 1 presents the descriptive statistics of all listed cars in the sample. The average listing time is 6.67 weeks and the average car age is 5.58 years. The average initial and last listing prices are $16,007 and $15,345.

In the used-car market, dealers have heterogenous capacity to hold inventory, so we expect that dealers' inventory processes are heterogenous. For each product-dealer combination, we calculate the mode, median, mean, and standard deviation of the weekly inventory over all 52 weeks. Table 2 summarizes some descriptive statistics across all 4,416 product-dealer observations included in the working sample, showing a substantial heterogeneity in the inventory across dealers and products. Figure B.2 displays the distributions of those statistics in greater details.

Table 2: Descriptive Statistics: Cross-sectional Heterogeneity of Inventory

	Mean	Median	SD
Mean weekly inventory at product-dealer level	8.59	5.71	9.79
Volatility (SD) in weekly inventory at product-dealer level	2.98	2.27	2.54

Note: An observation is the time-series mean or SD of a dealer's inventory of a product over 52 weeks. The sample includes 4,416 product-dealer combinations.

4.3 The Dynamics of Prices, Sales and Orders

In this subsection, we analyze how a dealer's decisions of list prices, sales and new orders during a week are affected by its inventory at the beginning of a week, or in other words, its inventory state.[11] We examine whether these effects are consistent with the predictions of our model.

First, we examine how a dealer's decisions of how many new orders to place and how fast to sell relate to its initial inventory. To be specific, we run the following regression, which is the reduced-form of the optimal retail and wholesale decisions of intermediaries established in Proposition 2, controlling for additional un-modeled factors that may be relevant in the data:

$$y_{k,f,t} = \beta_0 + \beta_1 x_{k,f,t} + \beta_2 X_{k,-f,t} + d_{k,f} + u_{k,f,t}, \tag{16}$$

where $y_{k,f,t}$ is the log of dealer f's new orders or listing removals of product k in week t, $x_{k,f,t}$ is the log of dealer f's inventory of product k at the beginning of week t, $X_{k,-f,t}$ is the log of the total inventory of product k of all other dealers except dealer f at the beginning of week t, $d_{k,f}$ is the product-dealer fixed effect, and $u_{k,f,t}$ is an idiosyncratic error. Inspired by Pesaran (2006), we use $X_{k,-f,t}$ to capture the common seasonal factor/shock to all dealers of product k.[12]

We make the standard sequential exogenous assumption,

$$Cov(x_{k,f,t}, u_{k,f,t+s}) = 0, \forall s \geq 0,$$

that is, the inventory at the beginning of a week is not correlated with the shock this week, nor future shocks.

Here, the product-dealer fixed effects $d_{k,f}$ can be correlated with the inventory $x_{k,f,t}$ for all t. Figure B.2 evidently shows a significant heterogeneity among dealers' inventory management decisions, suggesting the necessity of introducing a dealer fixed effect. To eliminate the fixed

[11] Unfortunately, our dataset does not contain information about wholesale prices.

[12] We also use time dummies to control for the common seasonal factor. The coefficient of the effect of inventory is similar. The result is upon request.

effects $d_{k,f}$, we take first difference of equation (16) to get:

$$\Delta y_{k,f,t} = \beta_1 \Delta x_{k,f,t} + \beta_2 \Delta X_{k,-f,t} + \Delta u_{k,f,t}. \tag{17}$$

The inventory at the beginning of a week $x_{k,f,t}$ is the last week's inventory plus the new orders and minus the sales of last week, where the last week's new orders and sales depend on last week's shocks $u_{k,f,t-1}$, so $x_{k,f,t}$ depends on $u_{k,f,t-1}$. As a result, $Cov(\Delta x_{k,f,t} \Delta u_{k,f,t}) \neq 0$, and hence the OLS estimation of the equation (17) can not give consistent estimates.

To deal with this problem, we use lagged inventory as an instrument. Next we argue that the inventories lagged at least two periods, $x_{k,f,t-s}$ for $s \geq 2$, are valid instruments for the change of inventory $\Delta x_{k,f,t}$. First, by the sequential exogenous assumption, $x_{k,f,t-s}$ for $s \geq 2$ are not correlated with current and last period shocks, and hence not correlated with $\Delta u_{k,f,t}$. Second, the current inventory is correlated with past inventory,

$$\begin{aligned} x_{k,f,t} = x_{k,f,t-1} + orders_{k,f,t-1} - sales_{k,f,t-1} = x_{k,f,t-2} + orders_{k,f,t-2} \\ -sales_{k,f,t-2} + orders_{k,f,t-1} - sales_{k,f,t-1} = ... \end{aligned} \tag{18}$$

The benefit of using lagged inventory for $s > 2$ is to minimize the possible serial correlation of the error term in the data. The top panel of Table B.1 reports the first stage results, where the dependent variable is $\Delta x_{k,f,t}$ and the instruments are the inventory lagged two weeks $x_{k,f,t-2}$ in the first column, the inventory lagged three weeks $x_{k,f,t-3}$ in the second column, the inventory lagged four weeks $x_{k,f,t-4}$ in the third column, and all the three lagged inventories in the last column. All estimates are statistically significant at 1% significance level, implying that the lagged inventories are correlated with $\Delta x_{k,f,t}$.

Table 3 reports the estimation results of the equation (17), where the dependent variable is the log of a dealer's new listings of a product in a week in the top panel and the log of a dealer's listing removals of a product in a week in the bottom panel. The first column reports the result of the OLS estimates and the other columns reports the results using the lagged inventory as instruments for $\Delta x_{k,f,t}$, where the second to forth column use the inventory lagged two weeks, three weeks, and four weeks as the instruments while the last column uses all of the three lagged inventories as instruments.

The estimates of the own inventory coefficient in the new-order equation (first panel of Table 3) are all negative and significant at 1% significance level, implying that a dealer tends to place fewer orders of a product in a week when it has a high inventory of that product at the beginning of that week. The estimates are smaller in magnitude when we use the lagged inventory as instruments. This suggests that the inventory is negatively correlated with the current shocks. The IV estimates

Table 3: Inventory v.s. New Orders and Listing Removals

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
(I) New Orders					
own inventory (β_1)	-1.6076***	-1.1688***	-1.1764***	-1.1636***	-1.1475***
	(0.0198)	(0.0371)	(0.0416)	(0.0461)	(0.0359)
aggregate inventory (β_2)	0.3694***	0.3752***	0.3750***	0.3734***	0.3737***
	(0.0056)	(0.0030)	(0.0030)	(0.0030)	(0.0030)
(II) Listing Removals					
own inventory (β_1)	1.3441***	1.3218***	1.3308***	1.3380***	1.3148***
	(0.0171)	(0.0367)	(0.0412)	(0.0455)	(0.0354)
aggregate inventory (β_2)	0.3617***	0.3612***	0.3611***	0.3595***	0.3592***
	(0.0055)	(0.0029)	(0.0030)	(0.0030)	(0.0030)
Instruments	-	$x_{k,f,t-2}$	$x_{k,f,t-3}$	$x_{k,f,t-4}$	$x_{k,f,t-s}$ for s=2,3,4
No. of Obs.	111,710	109,649	107,561	105,480	105,480

Note: Standard errors are in parentheses. $^{*}p < 0.10$. $^{**}p < 0.05$. $^{***}p < 0.01$.

are around -1.1 in magnitude, suggesting that a dealer places 11% fewer orders when the inventory is 10% higher. In other words, inventory order are roughly unit elastic to current stocks.

The estimates of the own inventory coefficient in the listing removal equation are all positive and significant at 1% significance level, implying that more listings of a product are removed in a week when the inventory of that product at the beginning of that week is high. The IV estimates are larger than the OLS estimates, suggesting that the inventory are negatively correlated with the current shocks. The magnitude of the IV estimates are around 1.3, implying that 13% more listings are removed when the inventory is 10% higher.

The estimates of the aggregate inventory are all positive and significant in both the new order and listing removals equations. This suggests that dealers are more active in placing orders and selling cars when the aggregate inventory (and likely demand) is high.

Next, we conduct a similar analysis for the listing price. The only difference is that the price is at the car-dealer-week level instead of the product-dealer-week level for the case of inventory orders and removals. We examine how a car's listing price relates to the dealer's inventory of that product at the beginning of that week, controlling for confounding factors including the log of weeks that car has been on sale, and the log of inventory of that product of all other dealers. To be specific, we examine the following equation, which is the reduced-form optimal retail pricing policy established in Proposition 2, controlling for additional un-modeled factors :

$$P_{j,k,f,t} = \gamma_0 + \gamma_1 x_{k,f,t} + \gamma_2 X_{k,-f,t} + \gamma_3 w_{j,k,f,t} + \mu_{j,k,f} + \epsilon_{j,k,f,t}, \tag{19}$$

where $P_{j,k,f,t}$ is the log of the listing price of car j of product k listed by dealer f in week t, $w_{j,k,f,t}$ is the log of weeks on sale, and $\mu_{j,k,f}$ is the car-dealer fixed effect. Again, to eliminate the car-dealer fixed effects $\mu_{j,k,f}$, we take first difference of equation (19) to get the following equation:

$$\Delta P_{j,k,f,t} = \gamma_1 \Delta x_{k,f,t} + \gamma_2 \Delta X_{k,-f,t} + \gamma_3 \Delta w_{j,k,f,t} + \Delta \epsilon_{j,k,f,t}. \tag{20}$$

Again, to deal with the correlation between $\Delta x_{k,f,t}$ and $\Delta \epsilon_{j,k,f,t}$, we use the lagged inventories as instruments. The bottom panel of Table B.1 reports the first stage results, showing that those lagged inventories are statistically correlated with $\Delta x_{k,f,t}$.

Table 4: Inventory and List Price

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
own inventory (γ_1)	0.0000	-0.0442***	-0.0501***	-0.0569***	-0.0296***
	(0.0003)	(0.0018)	(0.0023)	(0.0028)	(0.0019)
aggregate inventory (γ_2)	0.0006***	0.0002**	0.0002*	-0.0000	0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
weeks on sale (γ_3)	-0.0135***	-0.0048***	-0.0025***	-0.0122***	-0.0139***
	(0.0003)	(0.0004)	(0.0007)	(0.0009)	(0.0009)
Instruments	-	$x_{k,f,t-2}$	$x_{k,f,t-3}$	$x_{k,f,t-4}$	$x_{k,f,t-s}$ for s=2,3,4
No. of Obs.	1,197,254	997,538	833,894	698,216	694,964

Note: Standard errors are in parentheses. $^{*}p < 0.10$. $^{**}p < 0.05$. $^{***}p < 0.01$.

Table 4 reports the estimation results of equation (20), where the first column reports the OLS results and the other columns reports the IV results. All the IV estimates of the own inventory coefficient (γ_1) are negative and significant at 1% significance level, implying that a dealer tends to lower a car's listing price when its inventory of that product is high. The coefficient estimate in column (5) is around 0.03, suggesting that the listing price is 0.03% lower if the inventory is 1% higher. Given that the mean inventory is 21 and the mean listing price is \$16,335, this implies that one more car on list makes dealers to lower its price by \$23.

The estimates of the aggregate inventory coefficient are mostly significantly positive, suggesting that the listing price of used cars are higher when the inventory of the whole market is high, which is likely during periods of high demand. Also, the estimates of the time on sale are all negative, consistent with both the depreciation effect and the insight from the literature of observational learning.[13]

[13]See Taylor (1999), Kim (2017) and Kaya and Kim (2018) as examples.

4.4 A Closer Look at Inventories

Recall that our theoretical model predicts that there is a steady state ideal level of inventory for the intermediary. Whenever the actual inventory deviates from the ideal level, the intermediary works to return to the ideal inventory. As a result, the model predicts that the steady state distribution of inventory has a single mode (Proposition 3). Since we assume homogenous intermediaries and ignore the seasonality effect, the ideal level of inventory is time-invariant and constant across intermediaries. However, as displayed in Figures 1 and B.2, the inventory process exhibits both seasonality and idiosyncratic heterogeneity across dealers. They may significantly contribute to the shape of the inventory distribution. To control the seasonality and the product-dealer heterogeneity displayed, we normalize the inventory for each product-dealer combination to eliminate this heterogeneity.

Specifically, we assume that the ideal inventory $x^*_{k,f,t}$ of dealer f who sells product k at 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 homogenous and time-invariant, and the ideal 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 demean it:

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

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 ideal inventory level x^*_k for all product k dealers.

To examine the shape of the invariant distribution of the normalized inventory process, we

first need to test if the process is stationary. We use the Harris-Tzavalis unit-root test (Harris and Tzavalis 1999) for the normalized inventory shows that the panels are stationary with the estimated $\rho = 0.8762$ significant at 1% significance level. We also do the test for each product. The results show that the panels for each product are stationary. The estimated ρ for each product are all between 0.8 to 0.9 and significant at 1% significance level.

We test the unimodality of the inventory distribution for each product-dealer group by using the Dip Test (Hartigan and Hartigan 1985). All eight products pass the unimodality test at the 1% significant level, which are consistent with the prediction of Proposition 3. We display the distribution of normalized inventory for each product, pooled over all dealers and weeks, see Figure B.3 and Figure B.4.

Next, we test the unimodality of the inventory distribution for each product-dealer panel. Each panel has 52 observations. Among all 4,461 product-dealer panels, we cannot reject unimodality for 86% of the dealers. We further investigate the difference between the unimodal subsample and the non-unimodal subsample. We first look at the mean inventory of each panel over these 52 weeks. The mean of the unimodal subsample is 7.82 and the mean of the non-unimodal subsample is 13.35, implying that the non-unimodal panels are almost twice larger in terms of inventory. However, the two subsamples may sell different products and that is why their mean inventories are different. To examine this possibility, we take the difference between each panel's mean inventory over 52 weeks and the mean inventory of all dealers that selling the same product. The mean difference of the unimodal subsample is -0.62 and the mean difference of the non-unimodal subsample is 3.83. It suggests that even eliminating the product heterogeneity, the non-unimodal dealers are still significantly larger than unimodal dealers.

Furthermore, the property of single-modality is preserved after aggregation. Figure 2 displays the distribution of the normalized inventory pooled over products, dealers and weeks. As a robustness check, we consider an alternative normalizations of the inventory by double demean. Figure B.5 in the Appendix display the distribution of the alternative normalized inventory. Clearly, it is unimodal.

4.5 A Closer Look at Prices

It has been well documented that the retail price distribution is unimodaled in a variety of markets where intermediaries are present such as groceries (Lach 2002), retail products (Hosken and Reiffen 2004, Kaplan and Menzio 2015), mortgage rates (Woodward and Hall 2012) and illegal drugs (Reuter and Caulkins 2004) etc. In this subsection, we examine the shape of retail price dispersion in the used car market.

Idiosyncratic difference among cars within the same type contributes to differences in their

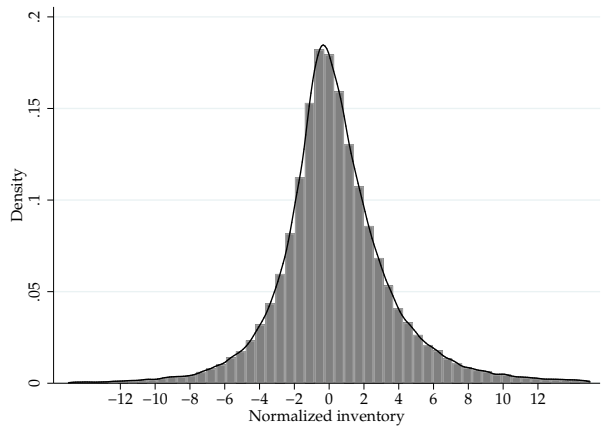


Figure 2: Normalized Inventory Pooled over Products, Dealers, and Weeks

Note: An observation in is a dealer's normalized inventory of a product at the beginning of a week, $\tilde{x}_{kft}$ defined in text. It includes 229,632 observations in total. The black solid line is the kernel fitting of the distribution.

prices, so first we need to net out these factors. Let P_{jkft} denote the log of the listing price of car j of product k holding by dealer f at time t, and w_{jkft} denote the log of the weeks that car j has been listed by dealer f on sale until time t. We run a regression of P_{jkft} on w_{jkft}, controlling for the random effects at the car-dealer level. Then, we take the mean of the price residuals within product (k), dealer (f), and week (t), denoted as p_{kft} to control the price difference caused by idiosyncratic difference among cars.

To control for the seasonality and the heterogeneity at the product-dealer level, we construct *normalized price* as

$$\tilde{p}_{kft} = p_{kft} - \bar{p}_{kf} - \bar{p}_{kt} + \bar{p}_k, \tag{23}$$

where $\bar{p}_{kf}$ is the average price residuals of all cars of product k listed by dealer f over all weeks, $\bar{p}_{kt}$ is the average price residuals of all cars of product k in week t across all dealers, and $\bar{p}_k$ is the average price residuals of all cars of product k across all dealers and over all weeks. We attribute the deviation of $\tilde{p}_{kft}$ from its mode to inventory dynamics.

We examine the unimodality of the distribution of the normalized price residual $\tilde{p}_{kft}$ for each product-dealer group using the Dip Test. We display the distribution of normalized price for each product, pooled over all dealers and weeks, see Figure B.7. They are all uni-modal. Next, we look at the distribution of price residual of each product-dealer combination, each with 52 observations. Among all 4,416 product-dealer groups, 99% are unimodal. Finally, Figure 3 displays the distribution of $\tilde{p}_{kft}$, pooled over products, dealers, and weeks. The dip test indicates that the

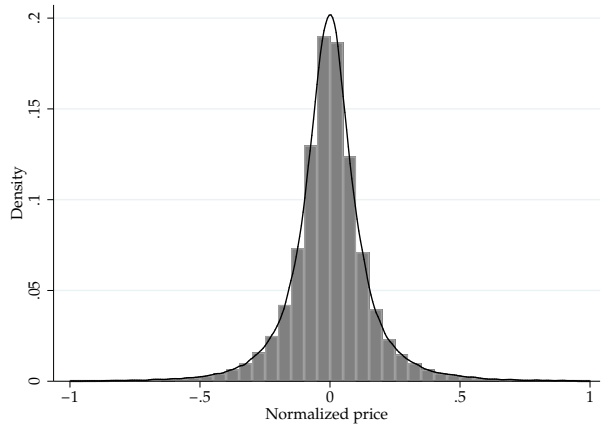


Figure 3: Distribution of Normalized Price Residuals

Note: An observation in is the normalized price of product k listed by dealer f in week t, $\tilde{p}_{kft}$ defined in text. It includes 229,632 product-dealer-week observations in total. The black solid line is the kernel fitting of the distribution.

normalized price is unimodal.

5 Concluding Remarks

This paper fills a gap between two active literatures: one on the role of intermediaries (Rubinstein and Wolinsky, 1987) and the other on pricing and inventory control (Whitin, 1955). We highlight the role of inventory dynamics on the shape of retail price dispersion and its dynamics. Most models of consumer search give rise to bi-modal price distributions, yet the empirical literature on price dispersion has documented unimodal distributions. We rationalize unimodal price distributions using a model where price dispersion is the result of intermediaries' optimal inventory and revenue management. Prices fluctuate as the intermediate finds itself away from the optimal inventory size, and intermediaries adjust prices to either sell invenotry or restock. We find support for this mechanism of price disperson in a large dataset on used cars sales.

A Appendix: Proofs

Proof of Lemma 1. To put it differently, we want to prove that

$$2V(x) \geq V(x-1) + V(x+1), \tag{24}$$

for every $x \in \mathbb{N}$. Since $\phi_r(\cdot), \phi_w(\cdot)$ are both strictly increasing and concave, their inverse functions $\phi_r^{-1}(\cdot), \phi_w^{-1}(\cdot)$ are well-defined, strictly increasing and convex. So the intermediary's problem (6) can be rewritten as

$$\begin{aligned} V(x) = -c(x) \quad &+ \quad \frac{1}{2} \max_{\rho_r \in [0,1]} \{\rho_r[u + \delta V(x-1) - \delta V(x)] - \kappa_b \phi_r^{-1}(\rho_r)\} \\ &+ \quad \frac{1}{2} \max_{\rho_w \in [0,1]} \{\rho_w \delta[V(x+1) - V(x)] - \kappa_s \phi_w^{-1}(\rho_w)\} + \delta V(x). \end{aligned} \tag{25}$$

where instead of choosing the market tightness, the intermediary directly chooses the probabilities of matching. Further define operator T as

$$\begin{aligned} TV(x) = -c(x) \quad &+ \quad \frac{1}{2} \max_{\rho_r \geq 0} \{\rho_r[u + \delta V(x-1) - \delta V(x)] - \kappa_b \phi_r^{-1}(\rho_r)\} \\ &+ \quad \frac{1}{2} \max_{\rho_w \geq 0} \{\rho_w \delta[V(x+1) - V(x)] - \kappa_s \phi_w^{-1}(\rho_w)\} + \delta V(x). \end{aligned} \tag{26}$$

It is easy to check that the standard Blackwell sufficient conditions hold, so the solution to problem (6) corresponds to the unique fixed point such that $V = TV$ in the set of continuous and bounded functions.

Fix an $x > 1$, denote $\rho_r^- = \rho_r(x-1), \rho_w^- = \rho_w(x-1), \rho_w^+ = \rho_w(x+1), \rho_w^+ = \rho_w(x+1)$, and let $\rho_r = \frac{\rho_r^- + \rho_r^+}{2}, \rho_w = \frac{\rho_w^- + \rho_w^+}{2}$. When the inventory is x, the policy (ρ_r, ρ_w) is feasible, so

$$\begin{aligned} 2TV(x) \quad \geq \quad & TV(x-1) + TV(x+1) + c(x-1) + c(x+1) - 2c(x) \\ & + \frac{\kappa_b}{2}\left[\phi_r^{-1}(\rho_r^+) + \phi_r^{-1}(\rho_r^-) - 2\phi_r^{-1}(\rho_r)\right] + \frac{\kappa_s}{2}\left[\phi_w^{-1}(\rho_w^+) + \phi_w^{-1}(\rho_w^-) - 2\phi_w^{-1}(\rho_w)\right] \\ & + \frac{\delta \rho_w^+}{2}[2V(x+1) - V(x+2) - V(x)] + \frac{\delta \rho_r^-}{2}[2V(x-1) - V(x) - V(x-2)] \\ & + \frac{\delta(2 - \rho_w^- - \rho_r^+)}{2}[2V(x) - V(x+1) - V(x-1)] \\ \geq \quad & TV(x-1) + TV(x+1). \end{aligned}$$

The first inequality holds because the policy (ρ_r, ρ_w) is feasible but may not be optimal. The second inequality results from the convexity of $c(\cdot)$, $\phi_r^{-1}(\cdot)$, and $\phi_w^{-1}(\cdot)$, and the "concavity" of

$V(\cdot)$ defined in (24), and the fact that ρ_w^- and ρ_r^+ are probabilities. Then by Corollary 1 of Theorem 4.1 in Stokey, Lucas, and Prescott (1989), we have the desired result.

□

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, an intermediary's inventory follows a Markov stochastic process $\{x_t\}$ determined by the equilibrium policy $\theta^*(\cdot)$ and $\lambda^*(\cdot)$. By standard argument (Theorem 2.2,2 in Ljungqvist and Sargent (2018)), 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

$$\phi_w(\lambda^*(x))g_{ss}(x) - \phi_w(\lambda^*(x-1))g_{ss}(x-1) = \phi_r(\theta^*(x+1))g_{ss}(x+1) - \phi_r(\theta^*(x))g_{ss}(x). \quad (27)$$

Because $\phi_w(\lambda^*(x))$ decreases in x, the left-hand side of (27) 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 (27) 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.

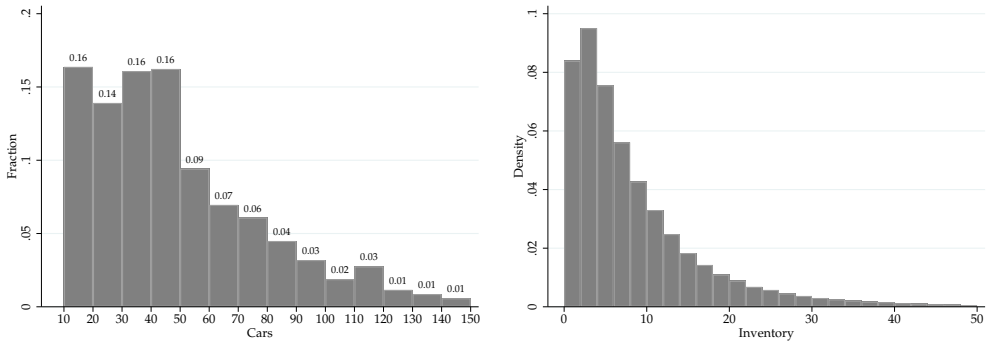
□

B Additional Tables and Figures

Table B.1: First Stage Results

The New Orders and Sales Equation (17)				
$x_{k,f,t-2}$	-0.1650***			-0.2035***
	(0.0014)			(0.0029)
$x_{k,f,t-3}$		-0.1264***		0.0103***
		(0.0013)		(0.0036)
$x_{k,f,t-4}$			-0.1024***	0.0364***
			(0.0013)	(0.0027)
$\Delta X_{k,-f,t}$	-0.0258***	-0.0222***	-0.0203***	-0.0262***
	(0.0008)	(0.0009)	(0.0009)	(0.0009)
No. of Obs.	109,649	107,561	105,480	105,480
The Listing Price Equation (20)				
$x_{k,f,t-2}$	-0.0429***			-0.1155***
	(0.0003)			(0.0012)
$x_{k,f,t-3}$		-0.0334***		0.146***
		(0.0003)		(0.0015)
$x_{k,f,t-4}$			-0.0275***	0.0640***
			(0.0003)	(0.0012)
$\Delta X_{k,-f,t}$	-0.0206***	-0.0188***	-0.0182***	-0.0208***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
$\Delta w_{j,k,f,t}$	0.0479***	0.0540***	0.0598***	0.0793***
	(0.0019)	(0.0028)	(0.0039)	(0.0039)
No. of Obs.	1,002,776	837,755	701,239	697,960

Note: Standard errors are in parentheses. $^{*}p < 0.10$. $^{**}p < 0.05$. $^{***}p < 0.01$.



(a) Dealer-Level Inventory (b) product-dealer-Week Level Inventory

Figure B.1: Inventory Distribution

Note: In Figure B.1a, an observation is a dealer's total inventory of all products, averaged over all weeks. It includes 713 observations in total. In Figure B.1b, an observation is a dealer's inventory of a product at the beginning of a week. It includes 229,632 observations in total.

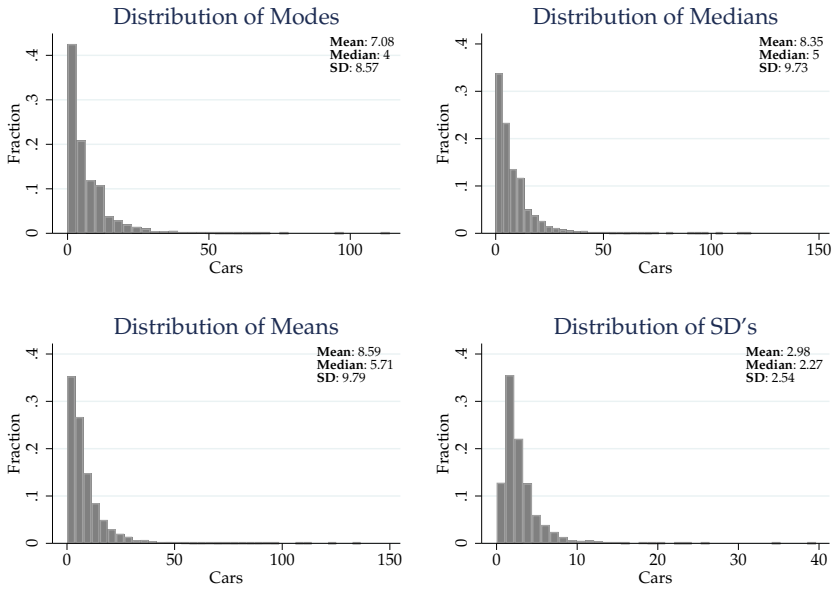


Figure B.2: Inventory Pooled over Products and Dealers: Mode, Median, Mean, SD

Note: An observation is the mode or median or mean or SD of a dealer's inventory of a product over 52 weeks. Each figure includes 4,416 observations.

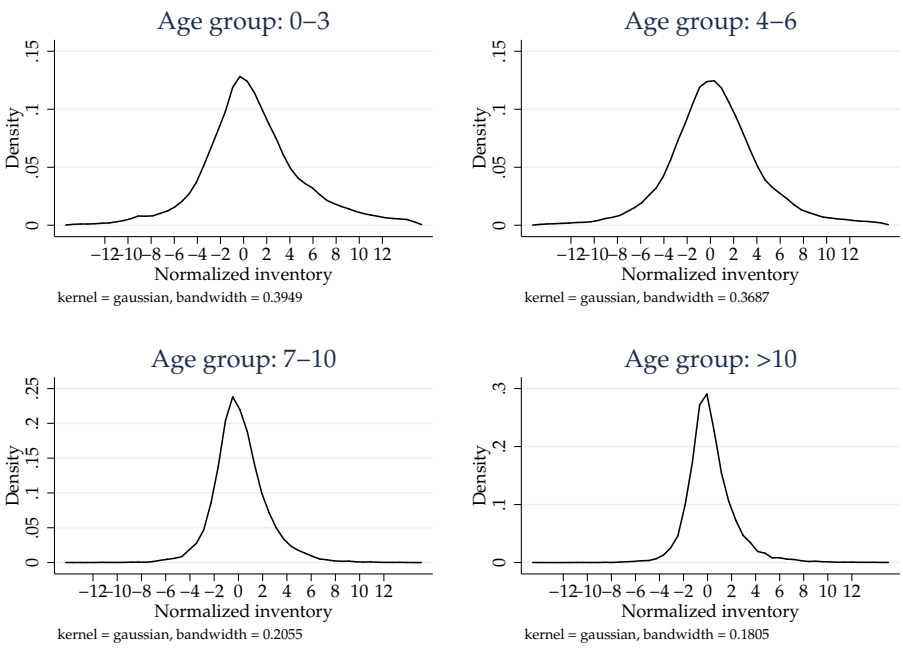


Figure B.3: Normalized Inventory Pooled over Dealers and Weeks: Sedan

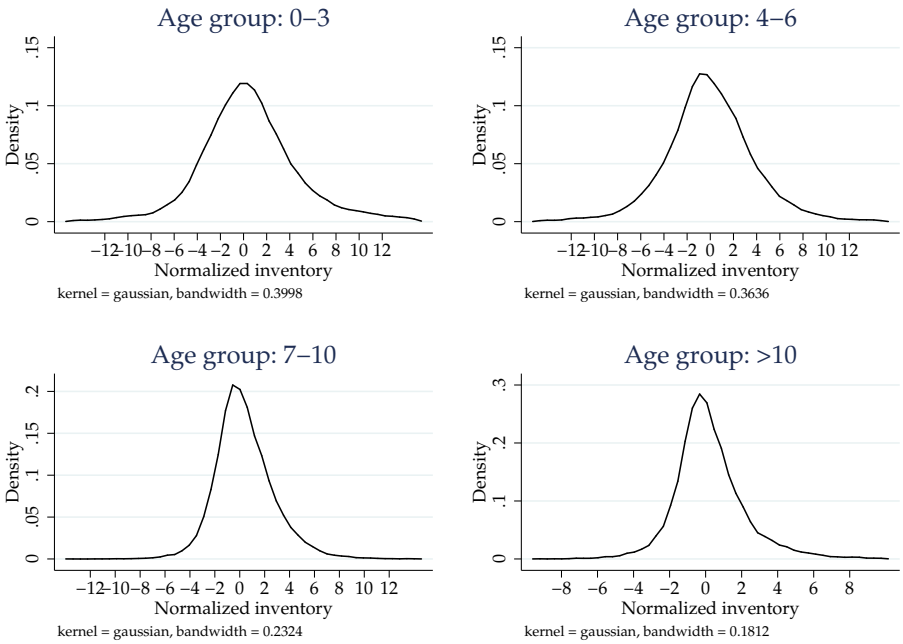


Figure B.4: Normalized Inventory Pooled over Dealers and Weeks: SUV

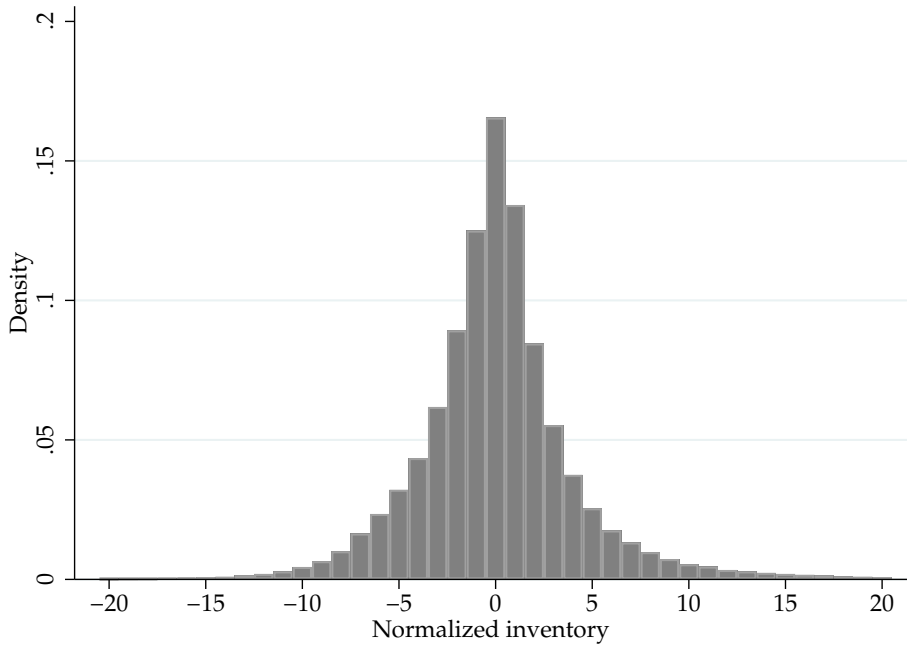


Figure B.5: An Alternative Normalizations of Inventory: Double De-mode

Note: An observation is the double de-mode of a dealer's inventory of a product in a week. It includes 229,632 observations.

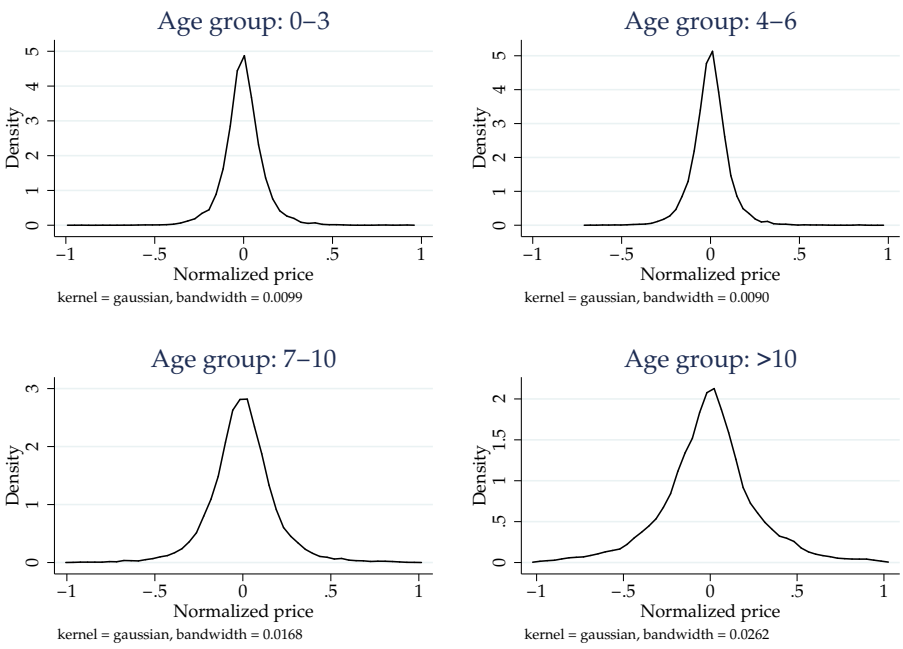


Figure B.6: Normalized Price Residuals of Sedan

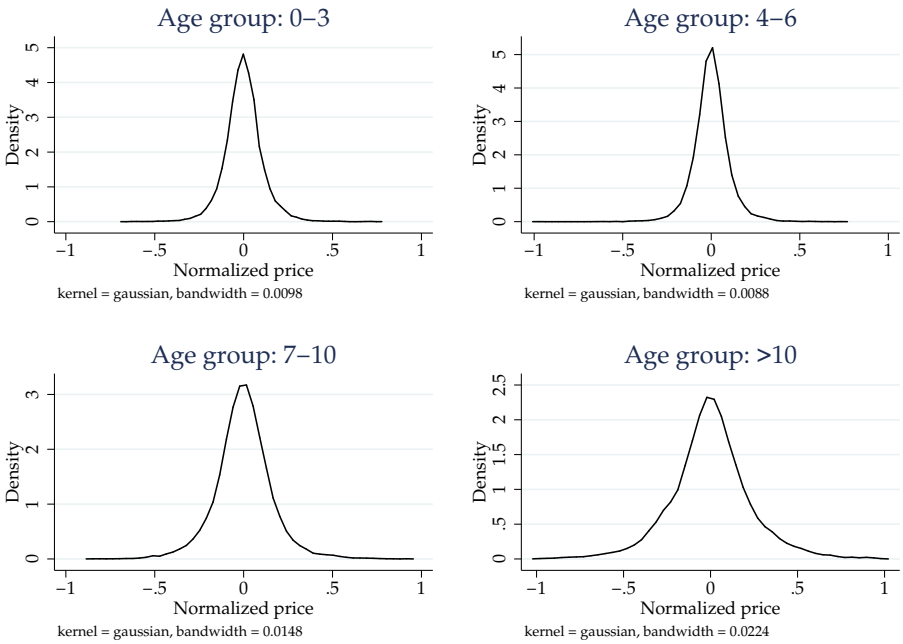


Figure B.7: Normalized Price Residuals of SUV

References

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

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 F. LI (2018): "Middlemen: the good, the bad, and the ugly," *The RAND Journal of Economics*, 49(1), 3–22.

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

BURDETT, K., AND D. T. MORTENSEN (1998): "Wage differentials, employer size, and unemployment," *International Economic Review*, 39, 257–273.

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

CACHON, G. P., S. GALLINO, AND M. OLIVARES (2018): "Does adding inventory increase sales? evidence of a scarcity effect in us automobile dealerships," *Management Science*.

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.

EDEN, B. (1990): "Marginal cost pricing when spot markets are complete," *Journal of Political Economy*, 98(6), 1293–1306.

FU, C. (2011): "Training, search and wage dispersion," *Review of Economic Dynamics*, 14(4), 650–666.

GARCIA, D., J. HONDA, AND M. JANSSEN (2017): "The double diamond paradox," *American Economic Journal: Microeconomics*, 9(3), 63–99.

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

GONZALEZ, F. M., AND S. SHI (2010): "An equilibrium theory of learning, search, and wages," *Econometrica*, 78(2), 509–537.

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.

HARRIS, R. D., AND E. TZAVALIS (1999): "Inference for unit roots in dynamic panels where the time dimension is fixed," *Journal of econometrics*, 91(2), 201–226.

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

HENDEL, I., A. NEVO, AND F. ORTALO-MAGNÉ (2009): "The relative performance of real estate marketing platforms: MLS versus FSBOMadison. com," *The American Economic Review*, 99(5), 1878–1898.

HOSKEN, D., AND D. REIFFEN (2004): "Patterns of retail price variation," *RAND Journal of Economics*, pp. 128–146.

JANSSEN, M., AND S. SHELEGIA (2015): "Consumer search and double marginalization," *American Economic Review*, 105(6), 1683–1710.

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.

KAPLAN, G., G. MENZIO, L. RUDANKO, AND N. TRACHTER (2019): "Relative price dispersion: Evidence and theory," *American Economic Journal: Microeconomics*, 11(3), 68–124.

KAYA, A., AND K. KIM (2018): "Trading dynamics with private buyer signals in the market for lemons," *The Review of Economic Studies*, 85(4), 2318–2352.

KIM, K. (2017): "Information about sellers' past behavior in the market for lemons," *Journal of Economic Theory*, 169, 365 – 399.

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

LIZZERI, A. (1999): "Information revelation and certification intermediaries," *The RAND Journal of Economics*, pp. 214–231.

LJUNGQVIST, L., AND T. J. SARGENT (2018): *Recursive macroeconomic theory*. MIT press.

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.

MOEN, E. R. (1997): "Competitive search equilibrium," *Journal of Political Economy*, 105(2), 385–411.

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.

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

PERLOFF, J. M., AND S. C. SALOP (1985): "Equilibrium with product differentiation," *The Review of Economic Studies*, 52(1), 107–120.

PESARAN, M. H. (2006): "Estimation and inference in large heterogeneous panels with a multi-factor error structure," *Econometrica*, 74(4), 967–1012.

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

REUTER, P., AND J. P. CAULKINS (2004): "Illegal 'lemons': price dispersion in cocaine and heroin markets," *Bulletin on Narcotics*, 56(1-2), 141–165.

RHODES, A., M. WATANABE, AND J. ZHOU (2018): "Multiproduct Intermediaries," Discussion paper, Yale University.

ROGERSON, R., R. SHIMER, AND R. WRIGHT (2005): "Search-theoretic models of the labor market: A survey," *Journal of economic literature*, 43(4), 959–988.

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

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*.

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.

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.

STOKEY, N. L., R. E. J. LUCAS, AND E. C. PRESCOTT (1989): *Recursive methods in economic dynamics*. Harvard University Press.

TALLURI, K. T., AND G. J. VAN RYZIN (2006): *The theory and practice of revenue management*, vol. 68. Springer Science & Business Media.

TAYLOR, C. R. (1999): "Time-on-the-market as a sign of quality," *The Review of Economic Studies*, 66(3), 555–578.

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

WATANABE, M. (2010): "A model of merchants," *Journal of Economic Theory*, 145(5), 1865–1889.

——— (2018): "Middle Men: The Visible Market-Makers," *The Japanese Economic Review*, 69(2), 156–170.

——— (2019): "Middlemen: A directed search equilibrium approach," Discussion paper, VU University Amsterdam, Tinbergen Institutey.

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.

WOODWARD, S. E., AND R. E. HALL (2012): "Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence," *American Economic Review*, 102(7), 3249–76.

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