Middlemen as Information Intermediaries: Evidence from Used Car Markets*

Gary Biglaiser Fei Li Charles Murry Yiyi Zhou†

October 22, 2017

Abstract

We theoretically and empirically examine used car dealers' roles as information intermediaries when asymmetric information is present. Our parsimonious theoretical model predicts that dealers' price premium (over private sellers) in dollar terms are hump shaped in car age, and in percentage terms are increasing in car age. It also predicts that dealer cars are more likely to be resold after transaction, due to their higher unobserved quality compared to cars sold by private sellers. We analyze rich datasets of universal used car transactions in two large U.S. states and our empirical findings are consistent with the theoretical predictions.

Keywords: Adverse Selection, Car Dealer, Information Intermediary, Used Car

JEL Classification Codes: D82, D83, L15, L62

*We thank Chao He, Qihong Liu, Alessandro Lizzeri, Brian McManus, John Rust, Henry S. Schneider, Randy Wright, Andy Yates, participants of the 8th Annual Madison Meeting on Money, Banking and Asset Markets, the 15th NYU IO day, and seminar participants at Stony Brook University for helpful comments.

†Gary Biglaiser: Economics Department, University of North Carolina at Chapel Hill, gbiglais@email.unc.edu. Fei Li: Economics Department, University of North Carolina at Chapel Hill, lifei@email.unc.edu. Charles Murry: Economics Department, Penn State University, cmurry@psu.edu. Yiyi Zhou: Economics Department, Stony Brook University, yiyi.zhou@stonybrook.edu.

1 Introduction

In modern economies, transactions are almost exclusively made through a variety of middlemen such as retailers, dealers, brokers, etc. Since there is no place for middlemen in a Arrow-Debrue's highly stylized world, to understand the ubiquitousness of middlemen, one must count on market frictions. One obvious rationale is offered by Rubinstein and Wolinsky (1987): middlemen can facilitate the searching and matching between trade parties in decentralized markets. Another popular justification of middlemen relies on frictions due to an *information asymmetry* between agents. As argued by Biglaiser (1993) and Lizzeri (1999), middlemen can serve as *information intermediaries* (or certifiers) in markets where there are selection issues. The idea is that middlemen have a more advanced technology and experience to distinguish product quality, so goods traded through them are of higher quality than those traded directly between sellers and buyers. The goals of this article are to theoretically and empirically examine the role of middlemen, car dealers in this case, in alleviating information asymmetry in the used car market.

A number of factors make the used car market suitable for our study. First, cars are complicated machines that require specialized care and maintenance; dating back to Akerlof (1970), the used car market has long been showcased as an example of a market rife with information asymmetries - sellers have more information about the product's quality than buyers do. Second, it is a large industry with retail sales totally over 500 billion dollars annually in the United States.[1] In 2016, 38.5 million vehicles were sold in the second-hand market in the U.S., more than twice the number sold in the new car market.[2]. Last, dealers are very active participants in the market. Nationally, about two-thirds of used car sales are made by dealers, and the other one-third occur between private parties. There are important differences between private sales and dealer sales. Private sales are much less regulated than dealer sales. Dealers are long run players who sell many cars and care about their reputations, while private sellers are in the market very infrequently and have little

[1] This number, constructed from Edmund's and Manheim yearly reports, represents revenues from franchised and independent dealers, so it is a conservative reflection of the size of the industry. We found conflicting reports about the total revenues of the private party sector.

[2] 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

reputation concerns. Furthermore, dealers can, and often do, provide a limited warranty to buyers. For example, CarMax, the largest used-car dealer in the U.S., provides a 15 days warranty and many other dealers provide similar short-term guarantee.

We build a parsimonious theoretical model to understand how a dealer's expertise to distinguish car quality is reflected in the market. We assume that as a car ages its value falls for two reasons. First, its quality, which is privately observed by the owner, may degrade. In the parlance of Akerlof, the car may become a lemon. Second, regardless of its quality, the value of the car depreciates over time due to use and changing consumer preferences. When faced with selling a car, the seller can visit a dealer who observes the private quality (by running a test), and the dealer decides how much, if anything, to offer for the car. The seller can either trade with the dealer or go to the market and sell the car directly to buyers. We assume the dealer receives a disutility by selling a lemon due to his reputation concerns. We show that, in equilibrium, the dealer trades with the seller if and only if the car quality is high. The market understands that the dealer is an expert and performs vehicle tests. Therefore, dealers trade higher quality cars on average and enjoy a price premium over the direct private sales.

The dealer's value as an information intermediation is reflected by the price premium it obtains relative to the market price. Furthermore, the vintage of cars has a considerable effect on the dealer's price premium for the following reasons. First, the information asymmetry between sellers and buyers develops as the car ages: Cars turn into lemons over time and seller become privately informed gradually. As a result, the dealer's value as an information intermediary grows as cars age. Second, the natural depreciation of cars reduce both the dealer's price and the market price and therefore the difference between these two prices falls when cars are sufficiently old. We show that the dealer's price premium in a dollar term is positive and hump-shaped due to the interaction of the natural depreciation of cars and the change in asymmetric information as cars age. When the dealer's price premium is computed as the ratio between the dealer's price and the market price, the depreciation effect appearing in both terms falls out. Thus, the dealer's price premium in percentage terms is monotonically increasing in a car's age.

The dealer's expertise of screening car quality generates a selection mechanism: cars purchased through dealers are more likely to be of high quality; while cars purchased directly from sellers are more likely to be lemons. By the classic adverse selection logic, agents who bought lemons will resell them sooner than high quality cars. As a result, our theory suggests that cars purchased directly from sellers are more likely to be resold in the near future than those purchased from dealers.

In summary, the theoretical framework generates two testable implications: (i) the dealer's price premium in a difference term is hump-shaped in the car's age; while the dealer's price premium in a percentage term is increasing in the car's age; (ii) the resale rate of a car is lower if it is purchased from a dealer than from the market.

We test the first implication by using a very rich data set of administrative records for used car transactions from the Virginia Department of Motor Vehicles. The data allow us to identify a make, model year, and exact trim with a particular set of options for all used cars traded from 2007 to 2014. Furthermore, we observe the transaction price, the odometer reading at the time of sale, whether the buyer and seller are a private party or a dealer, and in which county the transaction took place. By having such precise information about each car and transaction, we are able to difference out factors at the make-model-model year-trim level, the observable characteristics of a car, that may affect the price. We find that the correlations between a car's age and the dealer's price premium are consistent with the theoretical predictions of our model that we discussed above. Thus, we judge our results as strong evidence of adverse selection in the market and confirms the role of dealers as informational intermediaries.

To test the second implication, we use the data of universal used car transactions registered in Pennsylvania from 2014 to the middle of 2016, which allows us to keep track of each car's transaction history. We estimate a Logit model of whether a car is resold quickly after transaction on whether the car was purchased from a dealer, controlling for detailed car attributes and other factors. To deal with the concern that unobserved individual heterogeneity may correlate both with whether to buy a car from a dealer and with whether to resell it after transaction, we employ a two-step control function approach with the nearby dealers' inventory of the similar cars at the transaction time as the exclusive variable. We find that cars sold from private parties subsequently turn over faster, suggesting that they are of lower unobserved (unobserved by buyers) quality on average than dealer cars. We take this as strong evidence that adverse selection is present in this market. It also suggests that the dealer's price premium is not simply a placebo effect based on buyers' irrational belief of the dealer's expertise; instead, it is a reflection of buyers' rational expectation of dealers' role as information intermediaries.

We should state at the outset that there are other theories that are consistent with a positive dealer price premium over private seller transactions. In particular, car dealers as agents who can save on transaction costs due to search frictions or market segmentation due to dealer heterogeneity can explain these facts. On the other hand, these theories are not consistent with the humped shaped pattern in dollar terms of the price premium nor the increasing price premium in percentage terms that we find. We conclude that there still is a role for car dealers to save on transaction costs and as a agents to segment the market, but the role of them as certifiers must be a significant part of the analysis. We discuss alternative theories in more detail in Section 4.

1.1 Related Literature

Adverse Selection. Inspired by Akerlof (1970), economists have long studied whether information asymmetry exists in the leading example of a lemon market: the used car market. The evidence about adverse selection is mixed: Some find evidence of adverse selection; others do not; see Bond (1982), Bond (1984), Lacko (1986), Genesove (1993), and Engers, Hartman, and Stern (2009) as examples. Recently, inspired by the test derived by Hendel and Lizzeri (1999), Peterson and Schneider (2014) considers a car as an assemblage of parts, some with asymmetric information, and others without. They examine turnover and repair patterns and find evidence of adverse selection. We contribute to the literature by comparing the transaction quantity and price of dealers with those in direct sales. Rather than testing for the presence of asymmetric information by examining sellers' adverse selection, we focus on the selection made through dealers.

Middlemen. The theoretical foundations of this paper lie in the work of Biglaiser (1993), Biglaiser and Friedman (1994), and Biglaiser and Li (2017) who argued that in an environment with asymmetric information *a la* Akerlof (1970) middlemen emerge to identify lemons. There is also a literature discussing the function of middlemen as to agents to save search costs of agents in the market; see Rubinstein and Wolinsky (1987), Gehrig (1993), Yavaş (1996), Spulber (1996), Rust and Hall (2003), Wright and Wong (2014), and Nosal, Wong, and Wright (2015, 2017) as examples. Although a middleman's two aforementioned roles have been both well recognized on the theoretical side, the literature on the empirical side almost exclusively emphasizes that middlemen save search costs.[3] Gavazza (2016) shows

[3]One exception is Peterson and Schneider (2014), who report that cars sold by dealers require fewer repairs than cars sold by private sellers, although this is not their primary focus.

that dealers reduce trading frictions by correcting the misallocation of assets and lowering transaction prices. He also finds that the presence of dealers crowd out the number of agents' direct transactions. Recently, Salz (2017) investigated intermediaries' abilities to alleviate search costs in New York City's trade waste market. His results show that intermediaries improve welfare and benefit buyers in both the broker and the search market. Hendel, Nevo, and Ortalo-Magné (2009) compare house sales on a For-Sale-By-Owner (FSBO) online platform to the Multiple Listing Service (MLS) which only contains houses listed by realtors. They find that the price premium enjoyed by realtors is almost equal to the commission fee, but FSBO is less effective in terms of time to sell and the probability of a sale. Our paper contributes to this literature to empirically test a middleman's function as an information intermediary.

Quality Disclosure. There is a rich literature, both theoretical and empirical, on explicit quality disclosure statements by sellers. Grossman (1981) and Milgrom (1981) argued that the verifiable information disclosure by the seller himself can alleviate the information asymmetry. Lizzeri (1999) and Albano and Lizzeri (2001) study information disclosure by a third party certifier. Much of the work in the empirical literature focuses on the effect of explicit quality disclosure mechanisms, for example hygiene cards for restaurants in Jin and Leslie (2003), pictures accompanying eBay postings for cars in Lewis (2011), child care accreditation in Xiao (2010), and eBay's quality disclosure in Elfenbein, Fisman, and McManus (2015), Tadelis and Zettelmeyer (2015) and Cabral and Hortacsu (2010). In contrast, we analyze a mechanism that does not involve specific quality disclosure. Car dealers do not disclose that they sell high quality cars in a formal way. Instead, in equilibrium this is the correct expectations of consumers which is in turn reflected in a price premium. This mechanism for high quality in our model has a similar flavor to reputation, and is therefore closely related to empirical work in this area. For example, Jin and Leslie (2009) showed that chain restaurants have greater reputation incentives than independent establishments, but the implementation of an explicit rating system improves quality even further.

The rest of the paper is organized as follows. In section 2, we develop a theoretical model and state our two main testable implications with respect to car age effects and post-purchase resales. In section 3, we analyze detailed used car transaction data and present our empirical findings which are consistent with our theoretical predictions. In section 4 we offer and rule out some alternative explanations, and in section 5 we conclude and discuss directions for future research. We provide all proofs and additional theoretical analysis in Appendix A, and additional empirical analysis in Appendix B.

2 Theory

In this section, we derive predictions in a deliberately simple model where dealers play the role as information intermediaries. It illustrates how we expect dealers' roles as information intermediaries to affect the observed price and transaction patterns in the used car market.

2.1 Model

There is one used car. There is a seller, a dealer, and two (or more) buyers. We examine each car in isolation, since we treat the market demand by buyers as perfectly elastic at the car's expected quality.

Dynamics of Car Quality. The quality of the car is either high (H) or low (L). A car's age is $t \in [0, +\infty)$, and its quality changes over time by the following stochastic process: When new, $t = 0$, the car is of high quality. At each moment t, a quality shock arrives at a (failure) rate λ_t. Upon the arrival of the quality shock the car becomes low quality, $\theta_t = L$. We assume that low quality is an absorbing state.

Seller. The car is initially owned by the seller who privately observes the arrival of the quality shock. The seller remains passive until he receives a liquidity shock which arrives at a rate μ. A seller must sell his car upon the arrival of the liquidity shock.[4] The car's vintage, t, is *publicly observed*. The seller is able to visit (or get a price quote from) the dealer with probability $\alpha \in (0, 1)$ and goes to buyers directly if either he is unable to visit or does not make a transaction with the dealer. The α term is a reduced form modeling device which captures the probability that a seller cannot or decides not to sell through the dealer for some exogenous reasons. A seller's payoff equals the transaction price if he sells the car and zero, otherwise.

Buyers. There are at least two buyers. If a buyer pays p for a car of vintage t whose true quality is θ, his payoff is $U_t^\theta - p$, where U_t^θ represents the buyer's life-time payoff of owning a θ quality car vintage t. We normalize $U_t^L = 0$ and $U_t^H > 0, \forall t$. We assume that $\dot{U}_t^H \leq 0$ and

[4] We abuse the term of a liquidity shock to capture exogenous reasons for which the seller has to sell his car. Examples include the need to buy a new car, moving to other countries (states), etc.

$\lim_{t\to\infty} U_t^H = 0$, to capture the *depreciation effect*. That is, as the car ages, the marginal benefit of owning a high quality car rather than a low quality one is falling and eventually vanishes. In Appendix A, we argue that the depreciation effect can be obtained by assuming $\dot{\lambda}_t \geq 0$ and $\lim_{t\to\infty} \lambda_t = +\infty$. Buyers observe the car's vintage t, but do not know the current quality θ_t or whether the seller has visited the dealer. Buyers simultaneously bid for the good and observe the car's true quality once they take possession of it.

Dealer. The dealer observes the quality of the car and makes a take-it-or-leave-it price to the seller if the seller visits the dealer. If the dealer purchases the product at a price w and sells it to buyers at a price p, his payoff is

$$\begin{cases} p - w & \text{if } \theta = H, \\ p - w - k & \text{if } \theta = L, \end{cases}$$

where k captures the dealer's disutility due to selling a lemon. Such a disutility can be justified as a reputation loss or a monetary loss when he sells a car of low quality. We assume $k > U_0^H$ so that a dealer would not want to sell a lemon of any vintage.

Timing, Strategies and Equilibrium. At time 0, Nature randomly decides the arrival time of the liquidity and quality shocks for each seller. Although the quality of the car evolves over time, no player is making a decision before the arrival of the liquidity shock. Thus, we treat the arrival time t as a parameter and analyze the game upon the arrival of the liquidity shock at time t. We denote such a game by Γ_t which has four stages:

1. Nature decides whether a seller is able to find a dealer (with probability α).

2. If the seller finds a dealer, the dealer observes the quality of the car $\theta_t \in \{L, H\}$ and makes a take-it-or-leave-it offer, w, to the seller. Given the vintage t and the quality θ_t of his car, the seller then decides whether to accept the offer. If the dealer acquires the car, he sells it in the market where buyers bid simultaneously for the car and the dealer sells to the highest bidder.

3. If the seller fails to find the dealer or fails to trade with the dealer, he goes to the market. In the market, the buyers bid simultaneously for the car and the sellers sells to the highest bidder.

4. The wining buyer learns the true quality of the car immediately.

We solve the Perfect Bayesian Equilibrium (PBE): each player maximizes his or her expected payoff given his or her belief, and Bayes' rule is applied whenever possible. We focus on *trading equilibria* where the dealer trades car with positive probability.

2.2 Analysis

Since both the dealer and the buyers observe the car's vintage, t, denote q_t as the public prior belief that the car is high quality conditional on its vintage. Hence, by Bayes' rule, the process of $\{q_t\}_{t\geq 0}$ must obey the following differential equation:

$$\dot{q}_t = -\lambda_t q_t < 0, \forall t, \tag{1}$$

with the initial condition $q_0 = 1$. That is, the information asymmetry between the seller and buyers is developing over time: as the car ages, the public prior belief declines, with as $t \to \infty$, $q_t \to 0$.[5]

We analyze the game via backward induction. We begin with buyers' bidding behavior in the market. Because buyers bid as in Bertrand competition, $b_t = \hat{q}_t U_t^H$ where $\hat{q}_t$ denotes their equilibrium posterior belief conditional on the seller going to the market. In the equilibrium, the seller rationally anticipates his payoff b_t if he goes to the market, so he accepts the dealer's offer if and only if it is at least as attractive as b_t. Notice that $\hat{q}_t > 0, \forall t$ because $\alpha < 1$.

Now, we turn to the dealer's problem. The buyers' willingness to pay for a dealer's car is $\tilde{q}_t U_t^H$ where $\tilde{q}_t$ denotes their equilibrium posterior belief conditional on the car is traded through the dealer. Because $k > U_0^H$ and $\dot{U}_t^H \leq 0$, it is never optimal for the dealer to trade a lemon. In a trading equilibrium, the dealer purchases from the seller only if $\theta_t = H$, and the buyers bid U_t^H for the dealer's car. As a result, a high-quality car is traded in the private market only if the seller fails to find the dealer; and thus in the equilibrium,

$$b_t = \frac{(1-\alpha)q_t}{1-\alpha q_t} U_t^H. \tag{2}$$

The numerator is the measure of high-quality cars which are directly sold in the market and the denominator is the measure of all cars that are sold directly to buyers: those that never go to the dealer, $(1-\alpha)$, plus those that go to the dealer but are lemons who the dealer

[5]See Hwang (2016) for a more detailed discussion of developing asymmetric information.

does not buy, $\alpha(1-q_t)$. To maximize his profit, the dealer makes a *minimum winning offer* $w_t = b_t$ to the seller when $\theta_t = H$ and a *losing offer* $w < b_t$ when $\theta_t = L$. The former is the lowest offer that will be accepted by the seller; while the latter will be declined by the seller and results in a zero payoff to the dealer. Formally,

Proposition 1. *For any t, there is a unique trading equilibrium in which*

1. *In the private market, buyers bid the car's expected quality b_t satisfying (2).*
2. *The seller accepts the dealer's offer only if it is at least as large as his outside option, his expected market payoff, b_t.*
3. *The dealer makes a losing offer when $\theta_t = L$ and a minimum winning offer $w_t = b_t$ when $\theta_t = H$. The dealer sells the car at a price $p_t = U_t^H$.*[6]

In the equilibrium, the dealer trades with the seller only if $\theta_t = H$, causing an *adverse selection* effect on the set of the sellers going to the private market. Accordingly, the buyers will lower their belief of the quality of cars on the private market and thus their bids. The average quality of the cars traded through the dealer is U_t^H, which is higher than that of private sales, $\frac{(1-\alpha)q_t}{1-\alpha q_t} U_t^H$. The difference in the quality of cars traded through the dealer and those traded in the private market reflects two effects, one direct and one indirect. First, the dealer has a better technology to screen a high-quality car from a low-quality one thus an informational advantage. Second, since the dealer only purchases high-quality cars, the dealer's information advantage generates a adverse selection effect: it increases the proportion of low-quality cars in the market, which further enlarges the quality difference between the dealer's supply and the supply on the private market.

Fixing the car's vintage and other observable characteristics, we call the difference in the transaction price at the dealership and the market the dealer *price premium*. The dealer's price premium varies over the age of the car. Although both the dealer price, U_t^H, and the market price, $\frac{(1-\alpha)q_t}{1-\alpha q_t} U_t^H$, are decreasing in t, the driving forces for the declining price are different. The dealer's price declines simply because of the depreciation value of the car ($\dot{U}_t^H \leq 0$). On the other hand, the price of a direct transaction is decreasing because not only does the car depreciate but it is also more likely a lemon ($\dot{q}_t < 0$).

[6] There is another PBE where the dealer does not trade and buyers' beliefs about the dealer's car quality being high is sufficiently low off the equilibrium path. This equilibrium does not pass standard refinements such as intuitive criteria or D1.

First, we examine the age effect on the dealer's price premium in dollar terms:

$$p_t - b_t = \frac{1 - q_t}{1 - \alpha q_t} U_t^H. \tag{3}$$

To investigate the age effect, we take the derivative of (3) with respect to t and obtain

$$\underbrace{-\frac{(1-\alpha)(1-\alpha q_t)}{(1-\alpha q_t)^2} U_t^H \dot{q}_t}_{(+)} + \underbrace{\frac{1-q_t}{1-\alpha q_t} \dot{U}_t^H}_{(-)}. \tag{4}$$

The total age effect can be decomposed into two parts. First, it affects the dealer's value as an information intermediation. That is, it decreases the public prior belief q_t and thus the posterior belief of the buyers in the market, lowering the market price. Consequently, it increases the dealer's premium. This is captured by the first term of formula (4). Second, it decreases the buyer's willingness to pay for a high quality good, which is captured by the second term of formula (4). This is the standard depreciation effect. In general, the total effect of age on the premium can be *non-monotone*.

When $t = 0$, $q_t = 1$, so the second effect does not appear. Clearly, the price premium in (3) is strictly positive for $q_t < 1$, so the price premium in dollars is positive and increasing for small t. On the other hand, for very old cars, as $t \to \infty$, U_t^H goes to zero, and so does the price premium according to equation (3). Therefore, the price premium must eventually fall.

Second, one can also formalize the dealer's price premium over direct sales in percentage terms:

$$\frac{p_t}{b_t} = \frac{1/q_t - \alpha}{1 - \alpha}. \tag{5}$$

By taking the ratio between the dealer transaction price and direct transaction price, the depreciation effect, U_t^H, drops out and one can isolate the age effect through the change in q_t. That is, the change in the dealer's value of alleviating asymmetric information. Clearly, the formula in (5) is increasing in t.

Formally, we have our first empirical implication:

Implication 1. *The dealer's price premium in dollar terms formulated in (3) is positive for all car ages and is non-monotone in the car's age. For recent vintages it increases, and for sufficiently old cars it decreases: it is humped shaped. The dealer's price premium over* *direct sales in percentage terms formulated in (5) is greater than one for all car ages and is increasing in the car's age.*

An alternative way to understand the dynamics of premium premium is to compare the "declining rate" between direct sale price and dealer price as in Hendel and Lizzeri (1999). The direct sale price declines at a rate

$$\frac{\dot{b}_t}{b_t} = \frac{\dot{U}_t^H}{U_t^H} + \underbrace{\frac{\dot{q}_t}{q_t} + \frac{\alpha \dot{q}_t}{1 - \alpha q_t}}_{(-)},$$

which is faster than the price declining rate of the dealer price $\dot{p}_t/p_t = \dot{U}_t^H/U_t^H$. The declining of the dealer's price is driven by the depreciation effect only; while the declining of the market price also reflects the fact that older cars are more likely to be lemons.

2.3 Resales

Recall the classic logic of Akerlof (1970): asymmetric information causes cars that are observably identical to buyer sell at the same price even though they may actually be of different quality levels. Hence, owners of unobservably high quality cars sell less often because the seller reservation prices are higher than the market price. Our theoretical analysis predicts that dealer cars are of higher unobserved quality because dealers have informational advantages over buyers and they care about car quality due to reputation concerns. Therefore, we should expect that buyers of dealer cars are less likely to resell their cars because their cars are of higher average quality.

In this subsection, we extend our base model by allowing post-transaction resale. Recall that at stage 4 of our base model, the winning buyer immediately learns the quality of the car. We add a subsequent resale stage. At this stage, the winning buyer receives a liquidity shock with probability $\delta \in (0, 1)$ so that he has to sell his car in a resale market. We also allow the winning buyer to sell his car even if he does not experience a liquidity shock. The resale market observes the car's vintage, but can neither tell the buyer's motive for trying to sale the car nor whether the car was purchased from a dealer or directly from a private seller. As before, we assume the resale market is competitive and agents bid for the resale cars as in Bertrand competition. As $\delta > 0$, a high-quality car is resold with a positive probability, so the resale price $R_t > 0$. On the other hand, some low-quality cars will be resold too, so

$R_t < U_t^H$. Therefore, a high-quality car owner will resell his car only if he receives a liquidity shock, while a low-quality car owner will resell his car for sure.

If a buyer purchased the car from a dealer, he will resale the high quality car with probability δ. In contrast, if a buyer purchased the car from a seller directly, he will resell the car if either the liquidity shock arrives or if the car is a lemon. His resale rate in this case is given by

$$\frac{(1-\alpha)q_t\delta + (1-q_t)}{1-\alpha q_t}. \tag{6}$$

The numerator consists of sellers who sell their high quality cars directly to buyers who have a liquidity shock plus the measure of buyers who will sell their low quality cars that buyers want to sell, $(1-q_t)$. The denominator is the measure of all cars that are sold directly to buyers: those that never go to the dealer, $(1-\alpha)$, plus those that go to the dealer but are lemons who the dealer does not buy, $\alpha(1-q_t)$. Clearly, a car bought directly from a dealer has a resell rate greater than δ. Therefore, we derive another testable implication:

Implication 2. *A buyer is less likely to resell his car if the car was purchased from a dealer.*

In Appendix A, we show that when the probability of liquidity shock is sufficiently small, the prediction regarding the dynamics of price premium in Implication 1 remains.

2.4 Discussion

Before moving to the empirical analysis of the model implications, we discuss some of the model's features.

Asymmetric Information. The asymmetric information about the quality of the car is critical for our tests. Imagine instead that θ is public information. The depreciation effect will still push down the average transaction price of a car. However, a dealer's value as an information intermediary no longer exists. As a result, the price premium disappears. Furthermore, without asymmetric information the age effect on the price premium cannot be accounted for.

Dealers as Experts. Our tests rely on the credible selection through dealers. For this to be case, a dealer must have more experience or an advanced technology than ordinary consumers do not have to identify a lemon. For simplicity, we assume the dealer observes the quality of the car. Our results remain if we assume the dealer observe an informative but noisy signal about the quality. Imagine that the dealer runs a test whose outcome s is either good (G, for passing the test) or bad (B, for failing the test), and $\phi = \Pr(G|\theta_t = H) = \Pr(B|\theta_t = L) \in (0.5, 1)$. The dealer's purchase strategy becomes $w : \{B, G\} \to \mathbb{R}$. After the purchase, the dealer perfectly learn θ_t and decide whether to sell the car and the selling price. We can show that there is a unique equilibrium where the seller makes a minimum winning purchase offer when $s = G$ and a losing offer when $s = B$ and sells high-quality cars only. The buyers bid $b_t = \frac{(1-\alpha)q_t+\alpha q_t(1-\phi)}{1-\alpha\phi q_t-\alpha(1-\phi)(1-q_t)}U_t^H$ in the market. One can further verify that both Implications 1 and 2 still hold qualitatively.

Dealer's Selection. Our test relies on adverse selection through dealers. To incentivize a dealer to work as an honest gatekeeper, he must value the quality of the cars that he trades. In our model, we directly assume that selling a lemon will cost k dollars of expected profit for the dealer. This assumption captures the fact that dealers are less myopic than private sellers. This can be justified by thinking of a dealer as a long run player in an infinite horizon game. If the dealer sells a lemon, buyers would be able to detect that he cheated with sufficiently high probability, and the dealer will not be trusted in the future and the loss of future profits outweigh the short-run gains from selling lemons. On the other hand, a seller who has only one car cares little about their future reputation. The disutility of selling lemon can also be justified by the use of warranty. For example, CarMax offers 15 days warranties, and many manufacturers offer certified-pre-owned cars which promise free service for three years after purchase. See Biglaiser (1993) and Biglaiser and Friedman (1999) for more discussion about warranties and intermediaries.

Losing Offers Made by Dealers. The dealer's selection mechanism also relies on the fact that he may "refuse" to trade undesirable cars. For simplicity, we consider the dealer making losing offers as the only channel to implement the selection. In reality, there are many other mechanisms ensuring such selection mechanism. For example, low quality car owners anticipate unattractive offers from the dealers and therefore choose to visit the dealers with a lower probability. Also, the dealers can purchase cars failing test at a sufficiently low price and sell it to other dealers in wholesale used-auto auctions.[7]

Seller's Self-Selection The classic lemons market model *a la* Akerlof (1970) emphasizes the seller's self-adverse selection effect: a low-type seller is more likely to sell his car than a

[7]See a more detailed discussion on the wholesale automobile auctions in Genesove (1993) and Larsen (2014).

high-type seller due to his lower reservation value. This effect requires that both buyers and sellers to value quality. To focus on dealer selection, we assume the quality is irrelevant to the seller and ignore the seller's self-selection effect in Section 2.1. However, our insights remain when the seller's self-adverse selection is allowed. Take the following extension of our base model. The seller obtains flow utility u_θ from owning a car where $u_H = 1, u_L = 0$. Upon the arrival of the liquidity shock, the seller has to sell the car. Unlike our base model, the seller can choose to sell the car prior to the arrival of the liquidity shock. Once the seller decides to sell the car, the continuation game is identical to our base model. The dealer and buyers observe neither the arrival of the liquidity shock nor the arrival of the quality shock. As the informed seller can time his selling time, the model becomes a dynamic signaling game with dropout risk as in Dilmé and Li (2016). In this extension there exists an equilibrium where the seller sells the car upon the arrival of either the quality or liquidity shock. Therefore, given the car vintage t, the public prior belief is $q_t = \frac{\mu}{\mu+\lambda_t}$, which is decreasing in t as long as $\dot{\lambda}_t > 0$. It is easy to show that the equilibrium implications regarding the dealer's selection and the dynamics of price premium are qualitatively similar to our base model.

Dealer's Market Power. For simplicity, we assume the dealer is a monopolist. The idea is to allow the dealer to extract the value of his information advantage from trade so that he enjoys a price premium. In a setting with multiple dealers, our tests are still valid as long as dealers have some market power. It is worth mentioning that the variation of the market structure of the number of dealers have ambiguous effects on the price premium. When there are multiple dealers and demand is elastic, competition will push up their purchase price w and push down their selling price. The multi-dealer effect also reduces the price of direct transactions for two reasons. First, there is a standard competition effect. Second, there is a selection effect emphasized by our model. The buyers infer that the seller must have been "rejected" by multiple dealers and therefore lower their willingness to pay. As a result, the total effect on the dealer's price premium is unclear.

3 Empirical Analysis

In this section we use car registrations data to test the two implications derived in the theory section: (i) the dealer price premium in a dollar terms is hump shaped with respect to car age and the premium in percentage terms is increasing in car age; and (ii) cars purchased from dealers are less likely to be immediately resold than privately purchased cars.

3.1 Age Patterns of Dealer Price Premium

In this section we focus on Implication 1 regarding the age patterns of dealer price premium. We first introduce the data used for our analysis, then we present the empirical model to examine the relationship between dealer price premium and car age, and lastly we report and discuss our empirical results.

3.1.1 Used Car Registration Data from Virginia

The main data we use to test Implication 1 is obtained from the Virginia Department of Motor Vehicles (VA-DMV). It includes the universe of used car transactions registered in Virginia from January 1, 2007 to December 31, 2014. For each registration, we know the transaction date, price, the first 12 digits of the Vehicle Information Number (VIN) which is a unique number assigned to a vehicle that contains information to describe and identify the vehicle[8], and odometer mileage. We also know some information about the buyers and sellers. Sellers are either marked as "private seller," or as a dealer with a dealer identification number. We merge the dealer identification numbers with a separate dataset provided by the DMV that includes identification numbers matched to dealer names and addresses. Buyers are also marked as "private buyer" or with a dealer identification number. The zip code of buyers are also provided for many, but not all, observations. The zip codes of private sellers are also provided, but for many fewer transactions than for buyers.

Based on the information provided by edmunds.com, we decode the "squish VINs", the first 12 digits of VINs except for the ninth digit, into the make, model year, model, and exact trim with a particular set of options. The trim is a specific configuration of engine and other options available for a car. Most popular models have at least two trims available.

[8]The VIN standard, created by the National Highway Traffic Safety Administration (NHTSA) and enforced starting with the model year 1981, was required of all vehicle manufactured for use in the U.S. The NHTSA requires the VIN to be 17 digits long. The first three digits are reserved for the World Manufacturer Identification number and identify the manufacturer and country of origin of the vehicle. The fourth to eighth digits capture descriptive elements of the vehicle, including engine, body type, drive type, doors, restraint system and Gross Vehicle Weight (GVW) range. The ninth digit is a check digit that can be used to verify the validity of an encountered VIN using a calculation. The tenth digit identifies the model year of the vehicle and the eleventh digit identifies the specific plant and plant location that the vehicle was manufactured. The twelve to seventeen digits are serial numbers.

Table 1: Summary of Virginia DMV Data

	Mean	SD	Q25	Q50	Q75
Private Party Transactions					
Price	3,960	5,144	1,000	2,000	4,500
Mileage	134,376	67,290	92,183	132,315	171,300
Age	11.14	4.38	8	11	14
Dealer Transactions					
Price	13,032	8,518	6,349	12,000	17,779
Mileage	77,402	53,325	36,449	66,675	107,811
Age	5.99	4.05	3	5	9
Dealer Sales: 60.09%					
Total Observations: 5,469,241					

Note: The data includes all used car transactions in Virginia from January 1, 2007 to December 31, 2014. Sample selection is described in text. Data source: Virginia Department of Motor Vehicles.

For example, the squish VIN of 4T1BF3EKBU stands for 2011 Toyota Camry LE with a 4-cylinder engine and an automatic 6-speed transmission. Using the zip codes of buyers and sellers, we merge the DMV data with a list that matches zip codes to counties.

We make a number of sample selection decisions for the raw data in order to focus on our research questions. First, we drop 387,926 transactions when dealers are buyers.[9] Second, we discard transactions with negative odometer readings or when a car is more than 20 years old. We also discard transactions with recorded prices less than \$500 or greater than \$50,000. These transactions are outliers (for example transactions between family members) or were mistakenly recorded. In the end, our sample includes 5,469,241 transactions. Among them, 3,286,326 transactions (60 percent) were sold by car dealers and the remaining 2,566,349 (40 percent) were sold by private sellers.

In Table 1 we present the the sample statistics, including the transaction price, car age, and odometer mileage for the two segments. Overall, cars sold by dealers were substantially newer and more expensive than those sold by private sellers. Specifically, an average dealer car was around 6 years old and sold at a price of \$13,032, whereas an average non-dealer car was 11 years old and sold at a price of \$3,960. However, the standard deviations of car age

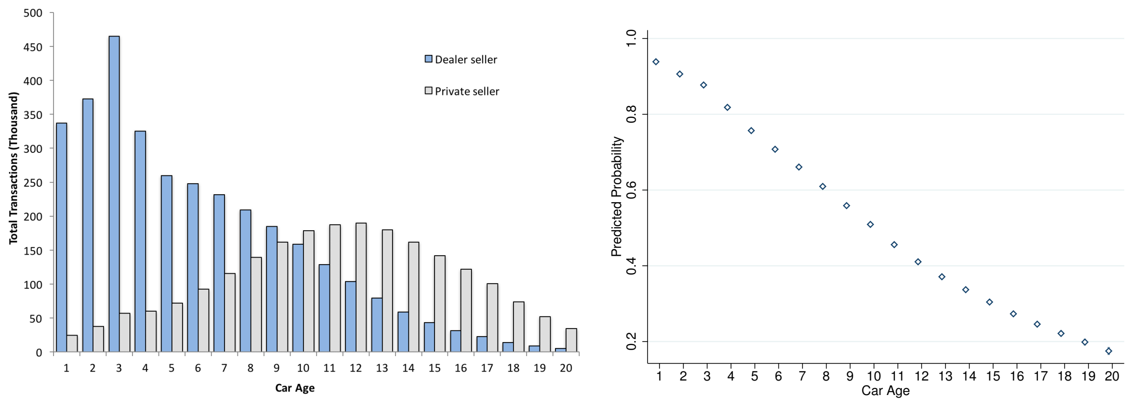
[9] Among them, 171,634 transactions were between dealers and 216,292 transactions were sold from individual sellers to dealers.

and transaction price are large, indicating that there were substantial heterogeneity across transactions.

In Figure 1a we present the total transactions of the two segments across different car vintages. First, the total number of dealer transactions falls in car age after peaking at three-year old cars, which is the common lease length for leasing cars. Second, the total number of transactions sold by private sellers increases in car age until age twelve and then falls in car age. To examine how the proportion of dealer sales relates to the car age, we estimate a linear probability model with *product* (make-model-model year-trim) fixed effects, where the dependent variable is an indicator for dealer seller and the key independent variables are a set of car age dummies. We also include other covariates that may be important predictors of the seller type: odometer mileage dummies for mileage in 30,000-mile intervals, monthly and yearly dummies, and dummies for seller county. All point estimates are statistically significant at least at the 0.001 level with standard errors clustered by product. In Figure 1b we plot the predicted probability that a car is sold by a dealer across car ages. Clearly, this probability falls in car age, from above 0.90 for relatively new vintages to below 0.10 for extremely old vintages. Recall that in our theoretical model, in equilibrium, the dealer trades with the seller only if the car is of high quality. Since the proportion of high quality cars declines in car age due to the natural depreciation of cars, the proportion of dealer sales falls in car age. Therefore, the declining pattern shown in the Figure 1b is consistent with the prediction of our theoretical model.

3.1.2 Dealer Price Premium and Car Age

Next, we examine how the dealer price premium relates to car age. To incorporate the heterogeneity across different cars in our data, we estimate a hedonic price regression where we regress log price on various transaction characteristics including car mileage, month and year effects, an indicator for dealer seller, indicators for different car ages, and age indicators interacted with the indicator of dealer seller. Importantly, we difference out any observed characteristics of cars by including *product* (make-model-model year-trim) fixed effects. The coefficients before the interaction terms of the dealer seller and car age indicators capture to what extent the dealer price premium co-varies with the car age. Essentially, we compare prices of two observationally equivalent cars (same model, same model year, same trim, same odometer mileage, and vintage), with one being sold a dealer and other one being sold by a private seller, and we examine how this price difference varies in the car age.



(a) Used-Car Transactions

(b) Predicted Probability of Dealer Sale

Figure 1: Dealer Sales and Private Sales

Note: An observation is a single used-car transaction in Virginia from 2007 to 2014. The predicted probability of dealer share is obtained from a linear probability model with product fixed effects. The dependent variable is one if the transaction took place at a car dealer and zero if the seller was a private party. Point estimates and predictions with 99% confidence intervals (not always visible due to size of marker) from the linear probability regression described in the text.

In specification (1), we include all used car transactions in our sample described above except for those extremely unpopular products with fewer than 100 transactions over the eight years (from 2007 to 2014) which account for less than 2% of the sample. We are left with 5,325,273 transactions, representing unique 35,248 model-model year-trims. In order to relieve the concern that new car dealers may take into account the substitution between their brand-new cars and used cars when they price their used cars, in specification (2) we limit our analysis to private sales and dealer sales from used-car-only dealers who do not have new car business lines. In addition, unpopular products may have liquidity issues which may affect their prices and induce correlation between search rents and car age. For example, if an older, desirable car has excess demand. To relieve this concern, in specification (3) we only include those most popular products that have more than 10,000 sales during the sample period. Lastly, to reduce the potential impacts of leasing cars, rental cars, and substitution from brand-new cars, in specification (4) we only include cars at least four years old.

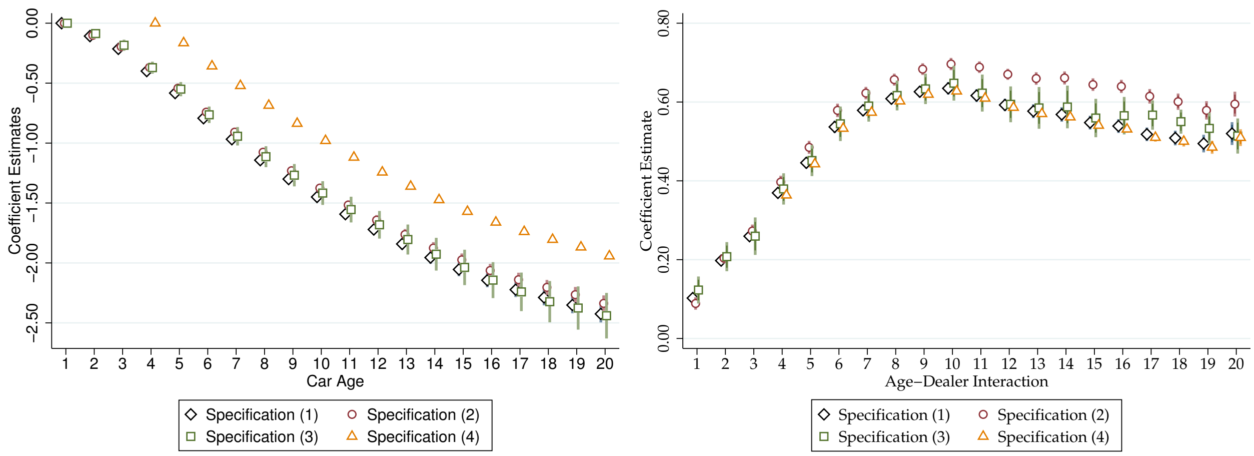
The estimation results are reported in Table 2 and Figure 2. The estimates are extremely precise, with every coefficient we report being statistically significant at least at the 0.001 level. As expected, the coefficient for the log of mileage is negative. The coefficients for car age indicators are reported in Figure 2a. Those coefficients are all negative and monotonically decreasing with age, implying that older cars are valued less. Notice that the age coefficients for specification (4) are above those for other three specifications. This is just because in specification (4) the baseline age is four year old rather than one year old in other specifications. Our primary focus is the coefficients for the age-dealer interactions, which are graphically reported in Figure 2b. The interaction coefficients are precisely estimated, and increase monotonically until age ten and thereafter level off and fall slightly.

Table 2: Dealer Premium Regressions

	(1)	(2)	(3)	(4)
log(Mileage)	-0.286	-0.326	-0.311	-0.375
Constant	12.553	12.904	12.736	13.098
Age Effects		*See Figure 2a*		
Age-Dealer Interactions		*See Figure 2b*		
R^2	0.498	0.471	0.547	0.460
Num. Observations	5,325,273	3,600,473	1,156,736	4,091,603

Note: An observation is a single transaction from the sample described in the text. The dependent variable is the log of transaction price and all specifications include product (make-model-model year-trim) fixed effects, log of the odometer mileage, month and year dummies, car age indicators, and interactions of age indicators and dealer seller indicator. All point estimates are statistically significant at least at the 0.001 level. Specification (1) includes the full sample. Specification (2) excludes cars sold by new car dealers. Specification (3) includes popular car models only. Specification (4) excludes cars younger than four years old.

Based on the estimates, we compute the predicted dealer premium in dollars across different car ages and display the results in Figure 3a. For all specifications the age profile of dealer premium is hump-shaped and reaches its peak at age six, at a value of between \$3,500 and \$4,000, depending on the specification. After age six the price premium declines monotonically until age twenty (less than \$1,000). Moreover, we compute the predicted dealer premium in percentage terms by car age and display the results in Figure 3b. The price ratio of dealer sales over private sales is increasing in car age until age ten, with a value of approximately 2 at that age, and then flattens and decreases slightly after age ten. It is not very surprising that our test loses power for older cars, since dealer sales dropped substantially for old cars, see Figure 1a. Overall, our results on the age patterns of the dealer price premium are consistent with the Implication 1 that we derive from the theoretical model.

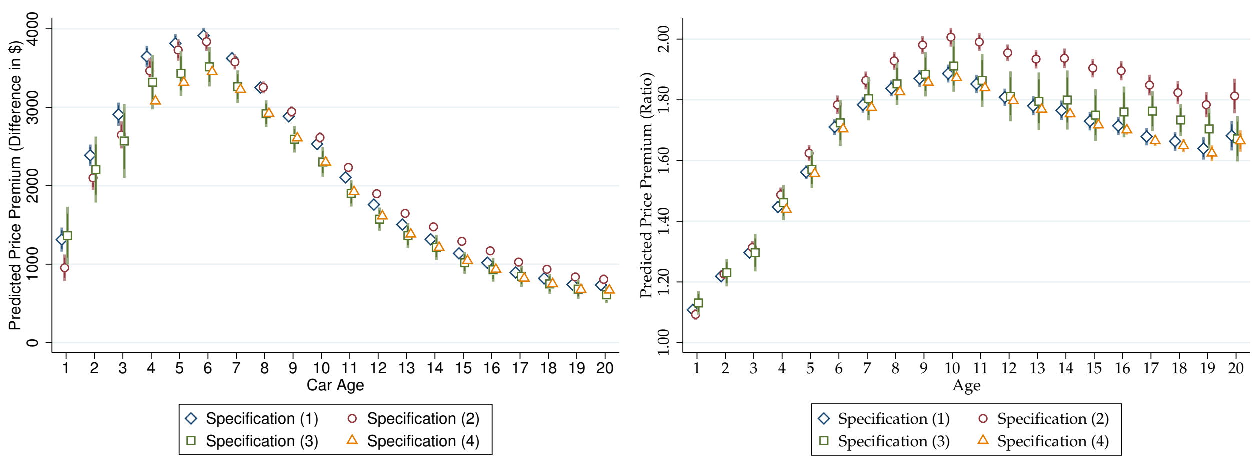


(a) Car Age Dummies

(b) Car Age-Dealer Interactions

Figure 2: Coefficient Estimates

Note: Point estimates with 99% confidence intervals. Different specifications refer to the different columns in Table 2.



(a) Price Difference

(b) Price Ratio

Figure 3: Predicted Dealer Premium

Note: Point estimates with 99% confidence intervals. Different specifications refer to the different columns in Table 2.

In addition, we estimate the four specifications by including seller county effects to control for those unobserved local factors affecting used car prices, and we present the predicted dealer price premiums across different car ages in Appendix B.1. The results are roughly the same as shown by Figure 3a and Figure 3b.

3.2 Post-Transaction Resales and Car Sources

In this section we test the implication 2 of our theoretical model that dealer cars are of higher unobserved (unobserved by buyers) quality than cars sold on the private market and as a result, dealer cars are less likely resold in the near future after transaction.

Our test of the relationship between resale rate and car source has been inspired by Akerlof (1970) and following empirical studies on testing the existence of adverse selection in the used car market, including Bond (1982), Bond (1984), Engers, Hartman, and Stern (2009), Peterson and Schneider (2014) and others. The key idea of the classic adverse selection tests is that owners of unobservably high quality cars sell less often. In our context, if dealer cars are of higher unobserved quality because of dealers' role as information intermediaries, then the turnover rate of dealer cars will be lower than that of cars sold by private sellers.

To ensure the validity of this test, the source of the car must be (i) the current owner's private information and (ii) observable to the researcher in the dataset. The first condition is more or less satisfied since the owner has no legal obligation to reveal the source of the car.[10] If this condition fails, cars bought from dealers and the ones bought from private parties should be considered being resold in different "markets" and the equilibrium market price will take the car source into account. To meet the second criteria, one must be able to trace the transaction history of cars. One limitation of our Virginia DMV data is that we do not observe the full VIN and as a result, we can not follow a car's transaction history. To deal with this issue, we obtain another dataset of used car registrations with the full VIN from the Pennsylvania Department of Transportation (PA-DOT).

3.2.1 Used Car Registration Data from Pennsylvania

The Pennsylvania data covers all used car transactions registered in this state from January 1, 2014 to July 31, 2016. The advantage of this dataset is that it includes the

[10] Carfax does not contain the type of previous transactions and may have unreliable information in general – see Murry and Schneider (2015).

Table 3: Resales after Purchase

	Dealer Sales		Direct Sales	
No. of Sales	648,106	(57%)	491,290	(43%)
Resale within one quarter	6,150	(0.95%)	10,865	(2.21%)
Resale within two quarters	12,938	(2.00%)	19,067	(3.88%)
Resale within three quarters	20,765	(3.20%)	27,183	(5.53%)
Resale within four quarters	29,661	(4.58%)	35,843	(7.30%)

Note: Percentage of cars transacted in 2014 resold after one, two, three, and four quarters.
Source: Pennsylvania Department of Transportation.

full VIN through which we can follow a car's post-transaction records. However, compared to the Virginia data, the price data is not as reliable and the time panel is substantially shorter. Therefore, we use the PA-DOT data only to examine buyers' reselling behavior after purchase.

The Pennsylvania data includes 1,861,473 used car transactions, with 57% of cars being sold by dealers and the remaining 43% being sold by private sellers. To study the relationship between the propensity of reselling and the car source, we focus on the transactions that occurred from January 1, 2014 till July 31, 2015, leaving the last year as a time window of post-purchase transactions. In the end, we have 1,139,396 unique cars that were transacted during this time period. Among them, 648,106 cars (around 57%) were sold by dealers. Table 3 reports the share of resales within different time windows, that is, one quarter, two quarters, three quarters, and four quarters, across different car sources (bought from dealers v.s. bought from private sellers). Regardless of the post-transaction time windows, the resale rates of dealer cars are substantially lower than those of cars sold by private sellers. For example, 0.95 percent of dealer cars were resold within one quarter after transaction, in contrast to 2.21% of cars sold directly by private sellers.

3.2.2 Logit Model with Product Fixed Effects

To further understand how a car's resale rate is affected by where it was bought from, we estimate a logit model with *product* (model-model year-trim-car age) fixed effects that control for cars' observable characteristics, analogous to our empirical strategy of the price regression:

$$y_i = 1\left\{\mu_i + \beta_d d_i + \boldsymbol{x}_i \boldsymbol{\beta}_{\boldsymbol{x}} + \epsilon_i > 0\right\} \tag{7}$$

Table 4: Immediate Resale after Purchase: Logit with Product Fixed Effects

	Resale Time Window			
	One Quarter	Two Quarters	Three Quarters	Four Quarters
Bought from Dealer	-0.392	-0.259	-0.186	-0.144
	(0.018)	(0.013)	(0.011)	(0.009)
Log Mileage	0.135	0.179	0.176	0.170
	(0.019)	(0.014)	(0.011)	(0.010)

Note: The dependent variable is an indicator for post-purchase resale within the specified time window. All specifications include model-model year-trim-car age fixed effects, monthly dummies, and county indicators. Standard errors in parentheses. The sample includes 1,139,342 unique used cars transacted from January 1, 2014 to July 31, 2015 in Pennsylvania. Sample selection is described in text. Source: Pennsylvania Department of Transportation.

where y_i indicates whether car i was resold within a specific time frame after transaction,μ_i are fixed effects at the model-model year-trim-car age level, d_i indicates whether the car was bought from a dealer, $\boldsymbol{x}_i$ is a vector, including the log of odometer mileage when the car was bought, monthly dummies, and indicators for the buyer's county to account for local differences in selling behavior, and $_i$ is an error term.

Table 4 report the estimation results of the Logit model for each of the four post-purchase resale time windows. Our primary coefficient of interest is the coefficient on whether a car was bought from a dealer (d_i). Our estimation results indicate that dealer cars are less likely to be resold for all four time windows we consider. Furthermore, this effect is decreasing in the number of quarters after purchase. This makes sense if defects are more likely to be discovered soon after purchase than later.

One concern is that some unobserved buyer characteristics may correlate both with where to purchase a car and with whether to resell it shortly after purchase. For example, purchasers who have high opportunity cost of time are more likely to buy cars from dealers and meanwhile, they are also less likely to resell their cars once they own them, implying a negative correlation between d_i and $_i$. Consequently, the Logit model in equation (7) without controlling for this unobserved buyer heterogeneity would over-estimate the impact of dealer seller on resale rate.

To deal with this potential endogeneity issue, we use a two-step control function, following a similar approach of Adams, Einav, and Levin (2009)'s analysis of delinquencies on subprime car loans. To do this, we need some variable that affects a buyer's choice of whether to buy from a dealer but does not directly affect her reselling decision. A reasonable candidate is the dealer inventory of cars with the same body type as the purchased product in the same week when the purchase occurred and in the same zip code of the buyer, denoted as z_i. The rationale is that a larger dealer inventory could provide buyers with more options and hence could attract more buyers to dealers. We obtained this information for transactions that occurred in four market areas in 2015 from cars.com. Our merged dataset includes 85,720 unique used cars transacted in those areas in 2015, along with their post-transaction records until the middle of 2016.

In the first stage, we run a Logit model of whether the car was originally purchased from a dealer on local dealer inventories (our exclusive variable) and other variables in the resale outcome equation. That is,

$$d_i = 1 \;\; \mu_i + \gamma_z z_i + \boldsymbol{x}_i \boldsymbol{\alpha}_{\boldsymbol{x}} + \upsilon_i > 0 \;\; . \tag{8}$$

We find that the estimate of α_z is positive and significant at 5 percent level, which is consistent with our expectation that a used car buyer is more likely to buy from a dealer if dealers in her neighborhood have a larger inventory of the car model she is interested. In the second stage, we include the residuals from the first-stage regression, denoted as $\hat{\upsilon}_i$, in our Logit regression of resales. That is,

$$y_i = 1 \;\; \mu_i + \gamma_d d_i + \boldsymbol{x}_i \boldsymbol{\gamma}_{\boldsymbol{x}} + \gamma_v \hat{\upsilon}_i + \omega_i > 0 \;\; . \tag{9}$$

Since we only have one year data, we consider two time windows: one quarter and two quarters after transaction. The estimation results are reported in Table 5. The first two columns are the results for the Logit model with model-model year-trim-car age fixed effects, described by equation (7), and the last two columns are the results for the control function approach, described by equations (8) and (9). Again, cars bought from dealers are less likely to be resold shortly after purchase, with the effect being stronger for the first quarter than two quarters. The control function approach appears to correct an attenuation bias in the data, as the estimates of the dealer seller coefficient are less negative.

Table 5: Immediate Resale after Purchase: Control Function

	Fixed Effects Logit		Control Function	
	Resale Window		Resale Window	
	One Quarter	Two Quarters	One Quarter	Two Quarters
Bought from Dealer	-0.540	-0.441	-0.488	-0.441
	(0.063)	(0.045)	(0.062)	(0.045)
Log Mileage	0.212	0.272	0.323	0.273
	(0.070)	(0.051)	(0.122)	(0.093)

Note: The dependent variable is an indicator for post-purchase resale within the specified time window. All specifications include model-model year-trim-car age fixed effects, year-month dummies, and county indicators. In the control function panel, we use dealer inventory as the exclusive variable for whether a car was bought from a dealer. Standard errors in parentheses. The sample includes 85,720 used cars transacted in four areas of Pennsylvania in 2015. Source: Pennsylvania Department of Transportation and Cars.com.

4 Alternative Explanations

Up to now, we have shown that our proposed theoretical model in which car dealers serve as information intermediaries can explain the age patterns of dealer price premia. Furthermore, we have also argued that our empirical analysis of the used car buyers' post-purchase resell decisions provides strong evidence that asymmetric information is present and dealers do sell higher quality cars when taking into account all observable information. In this section, we discuss alternative hypotheses focusing on search frictions, an owner's holding cost, and liquidity of the car market that may be able to explain the age patterns of price premia and resales activity. In each of the alternative hypotheses, the quality of the car is considered as public information. We argue that these alternative hypotheses do not explain *all* empirical patterns that we have documented. Therefore, we conclude that alleviating information asymmetry is one of the roles, among others, that car dealers are playing in the used car market.

4.1 Search Frictions

Since Rubinstein and Wolinsky (1987), there is a literature on intermediaries' matching roles to save agents' search costs. It is also possible that dealers in the used car market enjoy a price premium *only* by reducing agents' search expenses. In the following, we propose three alternative hypothesis based on search frictions. To fix ideas, we consider cars with different ages as different goods and therefore sold in different submarkets. In each submarket, (1) a monopoly dealer can frictionlessly trade with other parties, and (2) buyers and sellers with idiosyncratic (physical and opportunity) costs of search decide whether to go to dealers or search directly for each other.

Random Search. First, we assume the distribution of search cost of buyers are *identical* in all submarkets. In each submarket, a dealer's price premium in dollar terms must equal the expected search cost of buyers, which must be constant across submarkets due to the assumption of random search. Let us denote it as Δ. Denote p_t as the price in the private market of cars with age t, so the price of a dealer's car is $p_t + \Delta$. The dealer's price premium in percentage term is $1 + \Delta/p_t$, which is increasing if p_t is decreasing due to the depreciation effect. However, the dealer's price premium in dollar terms is Δ for all car ages t, which is inconsistent with the data.

Search and Sorting. It is possible that buyers with different characteristics may target cars with distinct characteristics, resulting in a sorting between buyers and cars. One may wonder whether the combination of search frictions and sorting theory can explain the empirical pattern. Suppose that the distribution of search cost of buyers vary in different submarkets. With a search cost saving dealer in the market, it must be the case that agents with higher search costs need dealer's service more, so they are more like to sell/buy through a dealer. In addition, in submarkets where buyers' search cost are high on average, dealers are able to charge higher price premia for buyers to be indifferent between the dealers and private sellers. This implies a *positive* correlation between the dealer's market share and the price premium. However, our empirical results demonstrate that the dealer's market share is monotonically decreasing in age (see Figure 1), while the dealer's price premium (in dollar terms) is initially increasing and then falling in car age. Furthermore, the dealer's price premium in percentage terms is increasing in car age; thus a negative correlation. In neither formulation of the price premium can one expect an unambiguous positive correlation between the dealer's price premium and market share. This contradiction implies that a dealer's value in the used car market cannot be rationalized by search frictions alone.

Search and Market Thickness. Cars can be viewed as assets whose value depends on both the flow payoff it generates and the resale value. Therefore, it is natural to believe

Table 6: Weeks on the Market before Sale

Car Age	1	2	3	4	5	6	7	8	9	10
Mean	7.39	7.15	7.09	6.88	6.27	6.03	5.85	5.58	5.19	5.02
SE	6.76	6.59	6.51	6.48	5.97	5.80	5.89	5.81	5.58	5.60
No. of Cars	20,293	19,148	19,081	11,240	8,175	6,095	8,515	7,520	6,616	5,854
Car Age	11	12	13	14	15	16	17	18	19	20
Mean	4.96	4.71	4.39	4.27	4.25	4.08	3.86	3.61	4.18	3.85
SE	5.72	5.64	5.54	5.72	5.73	5.31	5.53	4.54	5.17	5.30
No. of Cars	5,039	4,016	3,036	2,127	1,654	1,158	748	594	360	293

Note: This table reports the means and standard deviations of weeks before sale for dealer cars by car age. It also reports the number of dealer cars that are on sale by car age. The sample includes 131,567 used cars sold by dealers in four areas of Pennsylvania from January 2015 to July 2016. Source: Pennsylvania Department of Transportation and Cars.com.

the price of cars will be affected by their "liquidity".[11] As illustrated by Duffie, Gârleanu, and Pedersen (2005), and Gavazza (2016), when the trading technology exhibits increasing returns to scale in a frictional decentralized market, trading costs decrease with trading volume and assets with a thicker market are more "liquid", i.e. easier to trade. Applying this logic in our setting, the cars traded in a thicker submarket have higher liquidity values. If a dealer's main function is to overcome search frictions or to make cars more liquid, their price premium should be lower in a thicker submarket. Meanwhile, one should also expect that more liquid cars are traded more quickly. Therefore, this liquidity hypothesis predicts that the dealer price premium and time on the market are positively correlated. To empirically examine this correlation, we merge the PA-DOT data with the Cars.com data, and we get the information of how long a dealer car has sit on the dealer's slot before sale.[12] Table 6 reports the number of dealer cars by age, the means and the standard deviations of the time on market by car age. It indicates that the time on the market before sale is declining in car age. Combining with our empirical findings of the relationship between dealer price premium and car age, this result contradicts with the predictions of the liquidity hypothesis.

Therefore, none of these alternative hypothesis based on search frictions can explain the dynamics of price premium in Implication 1. In addition, one may also expect a theory based on search frictions and selection to explain the relationship between the resale rate

[11] We thank Alessandro Lizzeri for encouraging us consider this alternative explanation.

[12] Unfortunately, we can not get the data on how long a privately-transacted car is on the market before sale.

and car source. Intuitively, if a buyer has higher search/transaction/opportunity cost, he is more likely to purchase from the dealer and less likely to resell his car in the near future. However, as we illustrated in Table 5, such an unobservable heterogeneity of buyers can be controlled for by using the dealers' inventory as an instrumental variable. Our estimation results strongly suggests that the difference in resale rates is driven by adverse selection through the dealers rather than other unobservable heterogeneity among buyers.

4.2 Other Functions of Dealers

Car dealers provide multiple service such as selecting more popular models, reconditioning cars, and providing financing services. These services are also valued by the market and reflected in the dealers' price premium. Based on these functions of dealers, we discuss some alternative hypotheses.

Unobservable Advantage of Dealers' Cars. Recall that in our main data set, we only observe the first 12 digits of the VIN of a car which summarizes the basic manufacture information of the car, but it is insufficient to identify the car's color, navigation system, premium package, etc. One may wonder whether the dealer's price premium can be attributed to the characteristics of their cars which are observed by buyers but unobserved by econometricians. It is possible that a positive dealer price premium can be partially explained by offering cars with more popular colors and premium package (such as navigation system, leather and heated seats, etc.). However, such a story fails to explain the car age effect on the seller's price premium.

To fix ideas, consider a simple model where dealers' cars have premium packages and cars in private market do not. For a t age car, denote Δ_t as the average additional value of its premium package and denote p_t as its market price. Thus, $p_t + \Delta_t$ is the dealer's price. In reality, the value of premium package of the car depreciates as much as the rest of the car (such as engines); $\dot{\Delta}_t < 0$ captures the depreciation of the premium package, and $\dot{p}_t < 0$ captures the depreciation of the rest part. To explain the increasing price premium in percentage term $(\Delta_t/p_t + 1)$, one must have that

$$\frac{\dot{\Delta}_t p_t - \dot{p}_t \Delta_t}{p_t^2} > 0, \forall t,$$

which requires the depreciation of premium package to be slower than the rest part of the car. In addition, to explain the hump-shaped age effect on the dealer's price premium in difference term, Δ_t must be initially increasing and eventually decreasing in t, which seems unlikely in reality where the value of premium package should be decreasing in t.

Collusion between Buyers and Private Sellers. Virginia has a state sales tax on the purchase of used cars. This one-time vehicle "Sales and Use" tax is 3% of a vehicle's gross sales price and is charged whenever a vehicle changes ownership.[13] It is reasonable to believe that it is easier for private sellers to collude with private buyers and shade the sales price a bit to save on the taxes. However, just like in the search frictions theory, the collusion story has a hard time explaining the car age effects on the price premium and dealer market share.

Heterogenous Dealers and Market Segmentation. The used car market may be segmented where dealers specialize by automobile age and model. New car dealers focus on late model used cars, while exclusively used car dealers often buy and sell relatively older cars. As Genesove (1993) found, new car dealers typically have large market shares of the used car business, which can roughly explain the fact that dealer market share is falling in the car's age. However, it is hard to explain the difference in price premiums between exclusively used car dealers and new car dealers. More importantly, we estimated the relationship between the price premium and the car's age for new car dealers and used car dealers separately and plot them in Figure 3 as well. We do not observe a systematic difference.

Holding Cost and Competition. Car dealers have non-trivial holding costs for maintaining an inventory. Since newer cars are worth more than older cars, the interest costs are higher for holding the former cars. One would conjecture that dealers are willing to take a lower markup to sell newer cars more quickly due to holding cars. Similarly, the supply for late model used cars is relatively large due to the large number of returned lease cars. Competition will reduce both the dealer's price and the market price, leaving the effect on price premium ambiguous in general. Suppose the competition affects the dealer's price more than the private transaction price, one would observe a smaller price premium for newer cars. However, these theories cannot explain the hump-shaped price premium in dollar.

In summary, these alternative hypotheses (and a combination of them) do not explain all empirical patterns documented, especially the declining of price premium for sufficiently old cars and the relationship between resale rate and car source. Therefore, we conclude that

[13] During the titling process, the state takes either the 3% rate or \$35, depending on which one is higher. See https://www.carsdirect.com/car-pricing/how-to-calculate-virginia-car-tax

information asymmetry has considerable effects in the used car markets and one of important roles, among others, of car dealers is to alleviate the inefficiency due to the presence of asymmetric information.

5 Conclusion

We find evidence that used car dealers serve as information intermediaries in lemon markets: transaction patterns of used cars fit the theoretical predictions of a model with asymmetric information, and are inconsistent with various alternative explanations, for example based on search frictions. We also argue that the patterns in the data cannot be explained by dealer heterogeneity. Although there is a large literature that suggests middlemen contribute to market efficiency through alleviating search costs, our analysis suggests that dealers also contribute to market efficiency by alleviating information asymmetry.

A Appendix: Additional Theoretical Analysis

A.1 Micro-foundation of U_t^θ

The structure of U_t^θ is micro-founded as $U_t^\theta \equiv \mathbb{E}\left(\int_t^{+\infty} u_\tau^{\theta_\tau} d\tau\right)$ where $u_t^{\theta_t}$ is the buyer's follow payoff by owning a θ_t quality product whose age is t. Notice that $\{\theta_t\}$ evolves over t. We normalize $u_t^L = 0$ and $u_t^H = 1$ for all t. Hence, $U_t^L = 0$ and $U_t^H = \mathbb{E}\int_t^{\tilde{t}} 1 d\tau$ where the expectation is taken over the random time at when the quality turns to low: $\tilde{t} = \inf\{\tau|\tau \geq t, \theta_\tau = L\}$. That is, an owner enjoys one unit of flow payoff at every moment until the quality of the product turns low. Notice that, for simplicity, we assume that there is no discounting and the buyers' outside options are zero. Moreover,

$$U_t^H = 1dt + (1-\lambda_t dt)U_{t+dt}^H = dt + (1-\lambda_t dt)[U_t^H + \dot{U}_t^H dt + o(dt)]$$

Rearranging the above equation and taking $dt \to 0$ yields the Hamilton-Jacobe-Bellman (HJB) equation, $\dot{U}_t^H = \lambda_t U_t^H - 1$. Because $\dot{\lambda}_t \geq 0$, $U_t \leq \int_t^\infty e^{-\lambda_t}\lambda_t 1 ds = 1/\lambda_t$. Hence, U_t^H is (weakly) decreasing in t. We further assume that $\lim_{t\to\infty}\lambda_t = \infty$, so $\lim_{t\to\infty} U_t^H = 0$. To see this, suppose that $U_t^H \geq \epsilon, \forall t$ for some $\quad > 0$. As λ_t goes to infinity and continuous, there exists $\hat{t}$ such that $\quad < 1/\lambda_{\hat{t}}$. Then we have $U_t < 1/\lambda_{\hat{t}} < \quad$, which is a contradiction.

A.2 Resale

We solve the model backwardly. Due to the liquidity shock, both high-type and low-type car are resold with a positive probability in any equilibrium, so the resale price $R_t \in (0, U_t^H)$.Therefore, low-quality car winning buyer resells for sure and high-quality car winning buyer resells only if he receives the liquidity shock. As a result,

$$R_t = \frac{\delta q_t}{\delta q_t + (1-q_t)} U_t^H \in (0, U_t^H). \tag{10}$$

at stage 5. The analysis of the first 4 stages are similar except that buyers rational expect the possibility of resale. Therefore, the dealer's price becomes $(1-\delta)U_t^H + \delta R_t$. That is, a buyer rationally expects the dealer's car is of high quality for sure, but he also anticipate the arrival of a post-transaction liquidity shock. In that case, he has to resell the car and his payoff equals the resale market price R_t. Similarly, the seller's price becomes

$$b_t = \frac{(1-\alpha)q_t}{1-\alpha q_t}[(1-\delta)U_t^H + \delta R_t] + \frac{1-q_t}{1-\alpha q_t}R_t.$$

That is, the car sold in the market is of high quality with probability $\frac{(1-\alpha)q_t}{1-\alpha q_t}$. In this event, the winning buyer keeps the car and obtains a continuation payoff U_t^H unless the liquidity shock arrives. He resells the car if either the car is a lemon or the liquidity shock arrives. In that case, his payoff if R_t. It is straightforward to verify the rest of equilibrium strategies unchanged. The price premium in difference becomes

$$\frac{1-q_t}{1-\alpha q_t}(1-\delta)(U_t^H - R_t) = (1-\delta)\frac{1-q_t}{1-\alpha q_t}\frac{1-q_t}{\delta q_t + (1-q_t)}U_t^H \geq 0.$$

where $\frac{1-q_t}{1-\alpha q_t}\frac{1-q_t}{\delta q_t+(1-q_t)}$ is increasing in t and equals zero at $t = 0$. So the price premium in difference still increases in t when t is small and decreases in t when t is sufficiently large. The price premium in ratio is

$$\frac{(1-\delta)U_t^H + \delta R_t}{b_t}$$

which converges to the expression in (5) as $\delta \to 0$.Therefore, the price premium in percentage term is increasing in t as long as δ is sufficiently small.

B Appendix: Additional Empirical Analysis

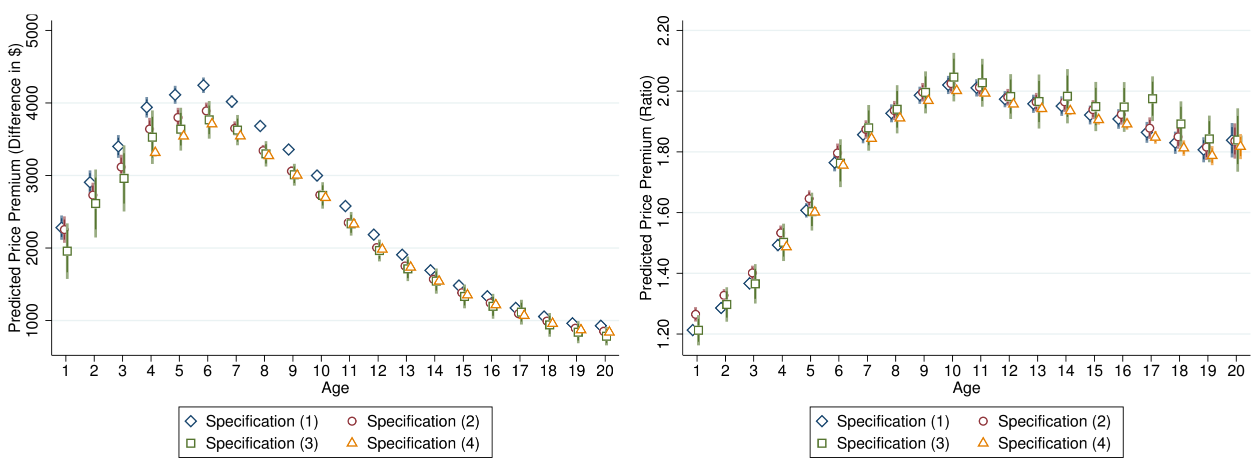
B.1 Price Regressions with Seller County Fixed Effects

Here, we report results from the price premium regressions and the consumer reliability regressions that include seller county fixed effects. One thing to note is that we do need observe seller county for about one million observations. We display the summary statistics from this reduced sample of transactions in Table 7. In Figure 4 we display the predicted price premium for price regressions that include dummies for the seller's county. The results are consistent with the baseline results in Figure 3.

Table 7: Summary of Virginia DMV Data, Seller Location Sample

	Mean	SD	Q25	Q50	Q75
Private Party Transactions					
Price	3,540	4,452	1,000	2,000	4,000
Mileage	137,590	65,031	96.720	135,421	174,245
Age	11.53	4.23	9	12	15
Dealer Transactions					
Price	13,314	8,317	6,990	12,500	17,900
Mileage	75,586	51,496	35,621	64,511	105,384
Age	5.92	4.00	3	5	9
Dealer Sales: 61.88%					
Total Observations: 4,147,299					

Note: The data includes all used car transactions in Virginia from January 1, 2007 to December 31, 2014. Sample selection is described in text. Data source: Virginia Department of Motor Vehicles.



(a) Dollar Term

(b) Ratio Term

Figure 4: Predicted Dealer Premium, With Seller County Fixed Effects

Note: Point estimates with 99% confidence intervals. Different specifications refer to the different columns in Table 2. Specification (1) includes 4,046,996 observations. Specification (2) excludes cars sold by new car dealers and includes 2,856,907 observations. Specification (3) includes only those popular products with 875,862 observations. Specification (4) includes cars older than three years old.

References

Adams, W., L. Einav, and J. Levin (2009): "Liquidity constraints and imperfect information in subprime lending," *The American Economic Review*, 99(1), 49–84.

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

Albano, G. L., and A. Lizzeri (2001): "Strategic certification and provision of quality," *International Economic Review*, 42(1), 267–283.

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

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

——— (1999): "Adverse selection with competitive inspection," *Journal of Economics & Management Strategy*, 8(1), 1–32.

Biglaiser, G., and F. Li (2017): "Middlemen: The Good, the Bad, and the Ugly," *Working Paper, University of North Carolina*.

Bond, E. W. (1982): "A Direct Test of the "Lemons" Model: The Market for Used Pickup Trucks," *American Economic Review*, 72, 836–840.

——— (1984): "Test of the Lemons Model: Reply," *American Economic Review*, 74, 801–804.

Cabral, L., and A. Hortacsu (2010): "The Dynamics of Seller Reputation: Evidence from eBay," *Journal of Industrial Economics*, 53, 54–78.

Dilmé, F., and F. Li (2016): "Dynamic Signaling with Dropout Risk," *American Economic Journal: Microeconomics*, 8(1), 57–82.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005): "Over-the-Counter Markets," *Econometrica*, 73(6), 1815–1847.

ELFENBEIN, D. W., R. FISMAN, AND B. MCMANUS (2015): "Market structure, reputation, and the value of quality certification," *American Economic Journal: Microeconomics*, 7(4), 83–108.

ENGERS, M., M. HARTMAN, AND S. STERN (2009): "Are Lemons Really Hot Potatoes?," *International Journal of Industrial Organization*, 27, 250–263.

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

GEHRIG, T. (1993): "Intermediation in search markets," *Journal of Economics & Management Strategy*, 2(1), 97–120.

GENESOVE, D. (1993): "Adverse selection in the wholesale used car market," *Journal of Political Economy*, pp. 644–665.

GROSSMAN, S. J. (1981): "The informational role of warranties and private disclosure about product quality," *The Journal of Law and Economics*, 24(3), 461–483.

HENDEL, I., AND A. LIZZERI (1999): "Adverse Selection in Durable Goods Markets," *American Economic Review*, 89(5), 1097–1115.

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.

HWANG, I. (2016): "Collapsing Confidence: Dynamic Trading with Developing Adverse Selection," *Working Paper, University of Miami*.

JIN, G. Z., AND P. LESLIE (2003): "The effect of information on product quality: Evidence from restaurant hygiene grade cards," *The Quarterly Journal of Economics*, 118(2), 409–451.

——— (2009): "Reputational incentives for restaurant hygiene," *American Economic Journal: Microeconomics*, 1(1), 237–267.

LACKO, J. M. (1986): "Product Quality and Information in the Used Car Market," *Bureau of Economics Staff Report to the Federal Trade Commission*.

Larsen, B. (2014): "The efficiency of real-world bargaining: Evidence from wholesale used-auto auctions," Discussion paper, National Bureau of Economic Research.

Lewis, G. (2011): "Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors," *American Economic Review*, 101, 1535–1546.

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

Milgrom, P. R. (1981): "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, pp. 380–391.

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.

Nosal, E., Y.-Y. Wong, and R. Wright (2015): "More on Middlemen: Equilibrium Entry and Efficiency in Intermediated Markets," *Journal of Money, Credit and Banking*, 47(S2), 7–37.

——— (2017): "Intermediation in Markets for Goods and Markets for Assets," *Working Paper, University of Wisconsin Madison*.

Peterson, J. R., and H. S. Schneider (2014): "Adverse selection in the used-car market: evidence from purchase and repair patterns in the Consumer Expenditure Survey," *The RAND Journal of Economics*, 45(1), 140–154.

Rubinstein, A., and A. Wolinsky (1987): "Middlemen," *The Quarterly Journal of Economics*, 102(3), 581–593.

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

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

Spulber, D. F. (1996): "Market Making by Price-Setting Firm," *The Review of Economic Studies*, 63(4), 559–580.

Tadelis, S., and F. Zettelmeyer (2015): “Information disclosure as a matching mechanism: Theory and evidence from a field experiment,” *The American Economic Review*, 105(2), 886–905.

Wright, R., and Y.-Y. Wong (2014): “Buyers, Sellers, And Middlemen: Variations On Search-Theoretic Themes,” *International Economic Review*, 55(2), 375–397.

Xiao, M. (2010): “Is quality accreditation effective? Evidence from the childcare market,” *international Journal of industrial organization*, 28(6), 708–721.

Yavaş, A. (1996): “Search and trading in intermediated markets,” *Journal of Economics & Management Strategy*, 5(2), 195–216.