MANAGEMENT SCIENCE
Vol. 61, No. 2, February 2015, pp. 431–453
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2014.1977

The Effect of Electronic Commerce on Geographic Purchasing Patterns and Price Dispersion

Eric Overby, Chris Forman
Scheller College of Business, Georgia Institute of Technology, Atlanta, Georgia 30308
{eric.overby@scheller.gatech.edu, chris.forman@scheller.gatech.edu}

The "law of one price" states that if prices for the same or highly similar goods vary across geographic locations by more than the cost of transport, then traders will shift supply and demand to exploit the price differences. However, several frictions prevent traders from doing this, including lack of information about prices and difficulty trading across locations. Electronic commerce has the potential to reduce these frictions by increasing price visibility and lowering transaction costs. We analyze this by studying how the diffusion of an electronic channel affected geographic trading patterns and price dispersion in the wholesale used vehicle market from 2003 to 2008. We find that buyers used the channel to shift their demand geographically to exploit price differences, which reduced geographic price dispersion. We find that the electronic channel also influenced how sellers distributed supply, but we find little evidence that this led to reduced geographic price dispersion.

Keywords: electronic commerce; electronic channels; electronic markets; price dispersion; geographic trade; wholesale automotive; auctions; law of one price; discrete choice; coarsened exact matching
History: Received October 27, 2009; accepted April 3, 2014, by Lorin Hitt, information systems. Published online in *Articles in Advance* September 22, 2014.

1. Introduction

The "law of one price" states that the price of the same or highly similar goods should not differ between any two locations by more than the cost of transport between them (including all relevant transaction costs). If the law of one price is violated, then theory suggests that buyers will exploit the price differences by shifting demand from high-price locations to low-price locations, and sellers will do likewise by shifting supply from low-price locations to high-price locations. However, several frictions prevent traders from shifting their supply/demand in this manner, including lack of information about prices and difficulty trading across locations (Baye et al. 2006, Stigler 1961). These frictions are particularly pronounced in geographically distributed markets and can segment a market into a series of regional submarkets across which prices may vary substantially. Electronic commerce has the potential to reduce these frictions by increasing price visibility and lowering transaction costs. Accordingly, we pose the following research question: how does the introduction of an electronic channel into a geographically distributed market influence buyer and seller trading behavior and geographic price dispersion? This question is important because excess geographic price dispersion indicates that goods are not being allocated efficiently, reflecting a type of market failure. Markets can improve the efficiency with which goods are exchanged, but their potential is greatest when information flows freely and transaction costs are low (McMillan 2002). Because new technologies (such as electronic channels) can improve information flow and reduce transaction costs, understanding their effect on geographic price dispersion is important for designing better markets.

We study this using transaction data from the U.S. wholesale used vehicle market from January 2003 to June 2008. This market is well suited to our analysis for two reasons. First, the wholesale used vehicle market has traditionally consisted of a set of segmented regional markets centered on market facilities located throughout the United States. Buyers, sellers, and vehicles are collocated at these facilities, where vehicles are sold via ascending auction in what we refer to as the physical channel. The market's geographic segmentation and resulting imbalances in supply and demand have caused prices for generally equivalent vehicles to vary across locations by more than the cost of transport. Second, a new electronic channel was implemented (in phases) during the sample period. This "webcast" channel streams via the Internet the live audio and video of the vehicles being auctioned at the physical market facilities and permits buyers to bid remotely. This reduces the frictions described above by allowing buyers to gather information about real-time prices at different facilities and by allowing them to purchase without having to travel.

The percentage of webcast purchases rose from approximately 0% to 20% over the sample period.

We find that buyers used the webcast channel to extend their purchasing reach to more geographically remote facilities, particularly when prices at those facilities were lower than those at facilities more geographically proximate. The shifting of demand from high-priced local facilities to low-priced remote facilities—which became more prevalent as the webcast channel diffused—lowered the geographic price dispersion in the market, which fell by approximately 16% from the beginning of 2003 to the beginning of 2008. We also find that the diffusion of the webcast channel influenced how sellers chose the facilities at which to sell vehicles, but little evidence that this was responsible for the observed reduction in geographic price dispersion.

Section 2 presents a brief literature review and the contributions of the study. Section 3 describes the wholesale used vehicle market and the data. Section 4 describes the nature of the geographic price dispersion in the market and the expected effect of the webcast channel. Section 5 presents our analysis of how the webcast channel influenced buyers' purchase behavior and how this, in turn, affected geographic price dispersion. Section 6 presents our analysis of seller behavior. Section 7 summarizes the study's findings and limitations.

2. Literature Review and Contributions

The paper draws on and contributes to two main research streams. The first is how electronic commerce affects geographic trade. A key theme in this stream is whether buyers use electronic channels to purchase from nearby or remote locations (e.g., Blum and Goldfarb 2006, Hortaçsu et al. 2009). The second is how electronic commerce affects price dispersion (e.g., Chellappa et al. 2011, Clemons et al. 2002). We link these two research streams by examining how buyers use an electronic channel to shift demand across geographic locations and how this affects price dispersion across those locations. We contribute to both research streams in several ways.

First, prior empirical studies have examined whether electronic channels lead to lower price dispersion (e.g., Aker 2010, Brynjolfsson and Smith 2000). Although many of these studies have demonstrated significant changes in price dispersion, they typically have not examined the microlevel behavioral mechanism that leads to that outcome. Observing this mechanism is critical for continued empirical research about electronic channels and price dispersion because different assumptions about the mechanism can result in more or less price dispersion when modeled analytically (Baye et al. 2006). As shown below, we use trend analysis, discrete choice methods, fixed effects panel regression, and matching estimation to examine the mechanism by which buyers' use of the webcast channel reduced geographic price dispersion. We examine buyers' use of both the physical and webcast channels, including how conditions at the physical market facilities influenced how buyers used the webcast channel. Our results show that buyers were more likely to shift purchases from nearby facilities where prices were relatively high to remote facilities where prices were relatively low when using the webcast channel than the physical channel. These demand shifts—which became more prevalent as the webcast channel diffused—are the mechanism that led to lower geographic price dispersion. A consequence of this finding is that as geographic price dispersion reduced, so should have the opportunities for spatial arbitrage in which an arbitrageur purchases a vehicle at facility A for price p and then quickly resells the same vehicle at facility B for a higher price p'. Indeed, Overby and Clarke (2012) found a negative correlation between buyers' use of the webcast channel and spatial arbitrage. However, similar to other studies in this general stream, they did not investigate the microlevel mechanism underlying this correlation; i.e., they did not empirically examine how and why use of the webcast channel led to reduced geographic price dispersion (and consequently to fewer spatial arbitrage opportunities).

Second, prior research on how electronic channels affect geographic trade has generally not focused on geographic price dispersion, and research on electronic channels and price dispersion has generally ignored the geographic location of products, although Jensen (2007) and Aker (2010) are exceptions (see below). The latter gap is because location has been irrelevant for the products that have often been studied (e.g., books, consumer electronics, and tickets) because the cost of shipping a product does not vary based on location and thus has minimal effect on its price. However, shipping costs vary significantly with location for products such as automobiles, agricultural commodities, fuels, and metals, the dollar value of trade for which is substantially larger than that for the products that have often been studied. These costs can have important effects on the locations from which buyers choose to purchase. We examine the location of products and buyers and find that the distance between them influenced how buyers used the physical versus the webcast channel and the facilities at which they purchased. This contributes new findings to the literature on how electronic commerce affects geographic trade.

Third, we examine how the diffusion of the webcast channel affected not only where buyers chose to purchase vehicles but also where sellers chose to sell vehicles. This differentiates our study from Aker (2010), who examined neither buyers' nor sellers' trading decisions, and Jensen (2007), who studied how the adoption of mobile phones affected where sellers sold fish but not where buyers bought them. Overby and Clarke (2012) investigated where sellers chose to sell vehicles in their analysis of how sellers' bounded rationality prevents them from making optimal distribution decisions for many of their vehicles (which subsequently creates opportunities to spatially arbitrage these vehicles). However, they did not examine the effect of the webcast channel on sellers' choices, nor did they examine buyers' choices. Among other findings, we show that sellers were more likely to distribute vehicles to facilities where the webcast channel was most widely implemented, likely because these facilities could attract more buyers. However, we find little evidence that this contributed to the reduction in geographic price dispersion.

3. Market Background and Data

3.1. Market Background

We study our research question in the context of the U.S. wholesale used vehicle market. This is a business-to-business market in which buyers and sellers trade used vehicles. The buyers are used car dealers who purchase vehicles in the wholesale market for resale to retail customers. Used car dealers procure approximately 30% of their used vehicle inventory via the wholesale market (NADA 2012, p. 11). The sellers are either other dealers or commercial sellers such as rental car companies (who sell vehicles retired from rental service) and leasing companies (who sell vehicles whose lease has expired). A dealer will sell vehicles in the wholesale market if he does not wish to (or cannot) sell them in the retail market. In that case, he will sell the vehicle wholesale to another dealer who will retail the vehicle. Commercial sellers sell in the wholesale market because they often lack retail operations and because the wholesale market is a highly liquid environment for selling multiple vehicles. Approximately nine million vehicles are exchanged in the market each year (NAAA 2013, p. 52).

3.1.1. Traditional (Physical) Market Operation. The market has traditionally operated as a physical market in which buyers, sellers, and vehicles are collocated at market facilities. These facilities are operated by intermediaries and are located throughout the United States as well as the world, although our analysis is specific to the United States. Sellers choose the facilities at which to sell their vehicles and then transport them there. Each facility has a large parking lot for vehicle storage and a warehouse-type building equipped with multiple *lanes*, which are essentially one-way streets. Transactions occur as follows. Managers at the facility determine the lane in which each vehicle will be auctioned. Each vehicle is driven—one at a time—down its assigned lane, in which buyers interested in that vehicle will stand. An auctioneer solicits bids from the buyers in an ascending, open outcry format. The vehicle is awarded to the highest bidder, if he meets the seller's reserve price, which is provided beforehand or indicated in real time by the seller (if she is present, which she typically—but not always—is). The vehicle is then driven away, and the next vehicle is driven into place and the process repeats. The auction for each vehicle takes approximately 30–45 seconds. Vehicles are usually auctioned in multiple lanes at the same facility concurrently. Transactions are conducted at least one day per week at each facility, sometimes more.

3.1.2. Electronic "Webcast" Channel. This physical process remains the predominant method by which vehicles are exchanged in the wholesale market in the United States. However, an increasing percentage of transactions are conducted electronically. The most popular electronic channel is the webcast channel. The webcast channel operates by streaming over the Internet live audio and video of the physical auctions occurring at the facilities. The webcast channel allows buyers to bid remotely—in real-time competition with the buyers who are physically present at the facility.

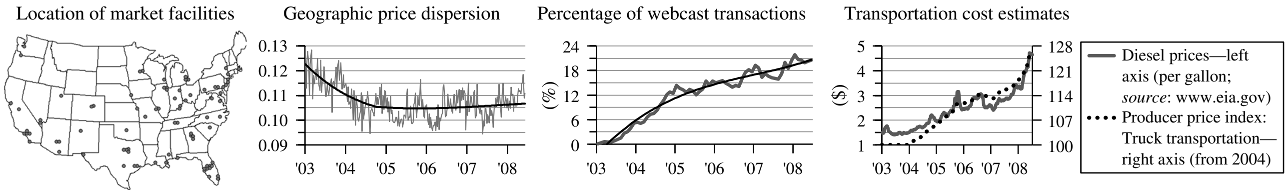
We highlight three points about the webcast channel. First, the webcast channel does not affect the basic auction process. This is because the auctioneer solicits bids for each vehicle in an ascending fashion, regardless of whether the bids are placed by collocated buyers using the physical channel or by "virtual" buyers using the webcast channel. Second, the webcast channel is specific to buyers; sellers do not use it. Sellers present their vehicles in the same fashion—having them driven down a lane at a physical market facility—regardless of whether buyers are using the physical or the webcast channels to place bids.[1] Third, the implementation of the webcast channel required that camera, microphone, and other equipment be installed *in each lane at each facility*. This proceeded in phases. During the early part of the sample time period, this meant that each facility had some lanes that were webcast enabled and some lanes that were not. We exploit this variance in the analysis reported in Appendix B.

3.2. Data

Data were provided by an intermediary in the wholesale used vehicle market that operates over 80 physical market facilities in the continental United States (see Figure 1). The data consist of all vehicles with between 15,000 and 21,000 miles that were auctioned (both successfully and unsuccessfully) at those facilities between January 2003 and June 2008. The mileage filter reduces

[1] The webcast channel is not the only electronic channel in the industry. There are also stand-alone electronic markets in the industry that operate similarly to eBay. We focus on the webcast channel, which was much more widely used than the stand-alone electronic markets during the sample period.

Figure 1 Location of Market Facilities and Temporal Trends by Week



Note. Polynomial trend lines shown.

Table 1 Variables and Descriptive Statistics

Variable	Description	Descriptive statistics
FacilityID	Denotes the market facility where the vehicle was auctioned.	There are 81 facility ID's.
FacilityZipCode	Zip code of each market facility.	—
LaneID	Denotes the lane at the market facility in which the vehicle was auctioned.	There are 12 lanes per facility (on average).
SellerID	Denotes the seller of each vehicle.	There are 28,791 seller IDs.
Sold?	Indicator variable for whether the vehicle was sold (1) or not (0).	Mean: 0.65
BuyerID	Denotes the buyer of each vehicle.	There are 74,917 buyer ID's.
BuyerZipCode	Zip code of each buyer.	—
Distance	Distance in miles between *FacilityZipCode* and *BuyerZipCode* for each sold vehicle.	Mean: 240.1; std. dev.: 339.1
Webcast?	Indicator variable for whether the vehicle was purchased by a buyer using the webcast (1) or physical (0) channels.	Mean: 0.12
ReceivedDate	Date each vehicle was received at the market facility.	—
AuctionDate	Date each vehicle was auctioned.	—
VehicleYear	Model year of each vehicle.	Mean: 2004.2, std. dev.: 2.5
VehicleMake	Make of each vehicle (e.g., Ford, Nissan).	There are 74 vehicle makes.
VehicleModel	Model of each vehicle (e.g., Ford Focus, Nissan Maxima).	There are 834 vehicle models.
Price	Sales price of each vehicle.	Mean: 14,852; std. dev.: 6,884
Valuation	Estimated market value of each vehicle. Calculated by the intermediary based on transactions for similar vehicles.	Mean: 15,017; std. dev.: 6,779
PriceValRatio	Price divided by *Valuation*.	Mean: 0.99; std. dev.: 0.09
Condition	Condition grade (0–5 scale, measured in 0.1 increments) assigned by the intermediary to indicate each vehicle's wear and tear. Higher grades indicate better condition.	Mean: 3.31; std. dev.: 0.70
Mileage	Odometer reading of each vehicle.	Mean: 18,069; std. dev.: 1,734
VehicleAge	Date each vehicle was auctioned minus *VehicleYear*. May be negative, e.g., if a 2008 model vehicle was auctioned in December 2007. Measured in years.	Mean: 1.48; std. dev.: 1.86

heterogeneity in vehicle condition, so that prices for vehicles of the same year and model across facilities may be compared. The low mileage of the vehicles in the sample also increases the likelihood that they are of predictable (and similar) quality, such that they may be purchased with confidence by buyers using either the physical or webcast channels (Overby and Jap 2009).[2] The data contain 3,588,975 auctions, 2,340,357 of which resulted in a purchase. Of these, 2,059,832 were purchased by buyers using the physical channel (88%) and 280,525 were purchased by buyers using the webcast channel (12%). The percentage of vehicles purchased via the webcast channel increased from just over 0% to approximately 20% over the sample period. We believe the sample is representative of the overall buyer population in the market because of the likelihood that most buyers purchased at least one vehicle with between 15,000 and 21,000 miles in the market between 2003 and 2008. Table 1 describes the data.

4. Geographic Price Dispersion and the Expected Effect of the Webcast Channel

Figure 1 illustrates the geographic price dispersion in the market and how it evolved during the sample period. We computed geographic price dispersion for Figure 1 as follows. First, we calculated the mean price of vehicles of year/model j (e.g., 2001 Toyota Camry) sold at each facility k in week t, referred to as the *facility means*. Second, we calculated the coefficient of variation of the facility means for year/model j in

[2] We determined the 15,000 to 21,000 mileage filter to be optimal after consultation with managers at the intermediary that provided the data. This filter balances two criteria: (a) that vehicles are lightly used and therefore have low quality uncertainty and (b) that there are enough vehicles in the sample to support econometric analysis.

week t (labeled CV_{jt}). For example, if the mean price for 2001 Toyota Camry's in the first week of May 2003 was \$13,500 at the Atlanta facility, \$12,000 at Orlando, and \$10,500 at Nashville, then $CV_{jt} = (1{,}500/12{,}000) = 0.13$. Third, we averaged CV_{jt} across all year/models j in each week to measure the overall dispersion of prices in the market per week (labeled CV_t). Because the coefficient of variation measures the variation of a variable relative to its mean, it is appropriate for comparing price dispersion across vehicle year/models (with different mean prices) and over time (Baye et al. 2006). Figure 1 also illustrates the growth in webcast purchasing. Geographic price dispersion declined by 16% between the beginning of 2003 and the beginning of 2008 (from 0.129 to 0.109) as webcast purchasing increased from 0% to 19%. The correlation between these time series for this period is -0.56 ($p < 0.01$).

The decline in geographic price dispersion was concentrated during the first half of the sample period, after which the trend is mostly flat with a slight increase. The flattening of the trend is deceptive for the following reason. The law of one price allows prices to vary across locations up to the cost of transport. Transport costs increased dramatically over the sample period due to the 209% increase in fuel prices (see Figure 1, far right panel). Thus, the baseline level of geographic price dispersion that is consistent with the law of one price increased over the sample period. As such, the flattening of the downward trend in the "nominal" geographic price dispersion shown in Figure 1 may represent a mixture of a downward trend in the "real" price dispersion with an upward trend in the baseline price dispersion. Indeed, geographic price dispersion reached a low of 0.094 in June 2007, just as fuel prices began their steep climb to record highs by mid-2008.[3] We explore this further in §5.4.[4]

4.1. Why Is There Geographic Price Dispersion?

The geographic price dispersion may exist if prices for vehicles of year/model j are consistently higher at some facilities than at others due to persistent imbalances in supply and demand across facilities. It may also exist if prices are *not* consistently higher at some facilities but instead vary week to week due to fluctuating imbalances in supply and demand.[5]

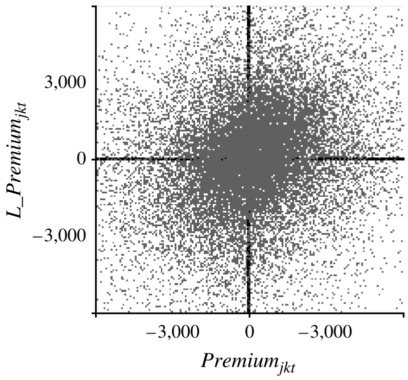
We examined which of these two possibilities was more likely by testing whether price premiums (or discounts) for vehicles of year/model j persisted over time at a facility. To do this, we measured $Premium_{jkt}$ as the mean price of vehicles of year/model j sold at facility k minus the mean price of vehicles of year/model j sold at all facilities in week t. ($Premium_{jkt}$ represents a discount when negative.) $L\_Premium_{jkt-}$ is similar but based on transactions *prior* to week t; we used lags of differing lengths for robustness. (Throughout, we use the $-$ subscript for variables based on events prior to time t.) First, we examined the correlation between $L\_Premium_{jkt-}$ and $Premium_{jkt}$. A correlation close to 1 (0) would indicate a high (low) degree of price premium persistence. Figure 2 shows a scatterplot of the $L\_Premium_{jkt-}$ and $Premium_{jkt}$ values (using a three-week lag for $L\_Premium_{jkt-}$), along with their correlation using different lag lengths for $L\_Premium_{jkt-}$. The correlation is small, approximately 0.13. Second, we computed the probability that $Premium_{jkt}$ is positive (negative) if $L\_Premium_{jkt-}$ was also positive (negative). A probability close to 1 (0.5) would indicate high (low) persistence. Depending on the length of the lag used for $L\_Premium_{jkt-}$, the probability ranges from 0.55 to 0.60. Third, we created a time series for each year/model j at facility k in which we entered a "+" if $Premium_{jkt} > 0$ and a "−" if $Premium_{jkt} < 0$ (we dropped instances in which $Premium_{jkt} = 0$, which occur less than 0.5% of the time). This generated sequences like "$++--+---++-+\ldots$" A "run" consists of uninterrupted +'s or –'s. A sequence with many runs such as "$-+--+-+-++----++$" would suggest low price persistence and a random fluctuation of above- and below-market prices, whereas a sequence with a low number of runs (e.g., "$++++++++-$") would suggest high price persistence and a nonrandom fluctuation. We tested this for each sequence using a runs test (Sheskin 2004). We could only reject the null hypothesis that the +'s and −'s occur randomly for 5% of the sequences, although the power of runs tests is low. Overall, there is little evidence of persistent price premiums (discounts) at facilities; the pattern is more consistent with regression to the mean than with persistent premiums. Thus, the geographic price dispersion is more likely to stem from imbalances in supply and demand that fluctuate from week to

[3] All results reported in the paper are qualitatively the same if we limit the sample period to 2003 to mid-2007.

[4] Both fuel prices and the percentage of webcast transactions increased over the sample period; the correlation between these time series is 0.84 ($p < 0.01$). Given this positive correlation and the negative correlation between webcast transactions and geographic price dispersion, the correlation between the fuel prices and geographic price dispersion time series is also negative ($\rho = -0.24$, $p < 0.01$). However, we show in §5.4 that the *conditional* correlation between fuel prices and geographic price dispersion—after accounting for changes in geographic trading attributable to webcast transactions—is positive, which is consistent with the law of one price.

[5] For example, consider scenario A in which average prices for 2003 Ford Rangers are \$9,000 in Phoenix and \$8,000 in Dallas each week, i.e., prices are consistently higher in Phoenix. Next, consider scenario B in which average prices are \$9,000 in Phoenix and \$8,000 in Dallas one week, the converse the next week, and so on. In both scenarios, the coefficient of variation of prices across Phoenix and Dallas is 0.083 each week.

Figure 2 Correlation Between $L\_Premium_{jkt-}$ and $Premium_{jkt}$ for Examining the Persistence of Price Premiums (Discounts) at the Same Facility



Length of lag for $L\_Premium_{jkt-}$	$L\_Premium_{jkt-}$ and $Premium_{jkt}$ correlation
1 week	0.11
3 weeks	0.13
8 weeks	0.13
13 weeks	0.14
26 weeks	0.14

week than from imbalances that persist over time.[6] Also, if price premiums (discounts) persisted over time, we would expect buyers and sellers to have recognized that well before the sample period and adjusted supply/demand in response (historical prices have been published by the intermediary since long before the introduction of the webcast channel). In that case, geographic price dispersion would be relatively flat throughout the sample period, rather than decreasing.

4.2. How Sellers and Buyers Might Exploit Geographic Price Dispersion and the Role of the Webcast Channel

Both buyers and sellers prefer facilities near them because vehicles are costly to transport and market participation has traditionally required physical attendance, at least for buyers. However, if geographic price dispersion is sufficiently large, then sellers may choose to distribute vehicles to remote facilities where they expect prices to be high. Similarly, buyers may choose to purchase vehicles at remote facilities where they expect prices to be low. If either sellers or buyers (or both) do this successfully, then supply and demand would become better balanced and geographic price dispersion would decrease. If price premiums (discounts) at a facility persist over time, then sellers and buyers could identify the high- and low-priced facilities and shift supply and demand accordingly to exploit the price differences. However, the above analysis shows that price premiums (discounts) do not persist very much; instead, there is significant fluctuation in where prices are high (or low) any given week. As such, the most effective way for sellers and buyers to exploit geographic price differences is to shift supply and demand in response to real-time price information at different facilities. As we discuss below, the webcast channel makes this feasible for buyers but not for sellers.

First, sellers distribute vehicles to facilities in advance, typically weeks before they are auctioned. The median difference between when a vehicle is distributed to a facility (*ReceivedDate*) and when it is auctioned (*AuctionDate*; see Table 1) is 18 days (mean = 27, s.d. = 36.6).[7] Thus, sellers who seek to exploit geographic price differences must forecast prices at different facilities weeks into the future, using information that will often be outdated by the time their vehicles are auctioned. To complicate matters, the obvious method of forecasting future prices based on recent prices appears to be only marginally effective at best (as shown above). The webcast channel does nothing to mitigate this for sellers because it does not affect the timing of when sellers distribute vehicles to facilities.

Second, buyers who use the physical channel must commit to a facility by traveling there. Only after these buyers have committed to a facility do they observe actual prices. If a buyer chooses a facility where prices turn out to be high, it is difficult for him to switch to another facility where prices might be lower. This is because the buyer would have to pay the real and opportunity costs of traveling to the other facility (either that day or another day). Thus, buyers who seek to exploit geographic price differences have traditionally needed to forecast future prices at different facilities, and they face similar challenges with this task as do sellers. *A key feature of the webcast channel* is that buyers do not have to commit to a facility. Instead, a buyer can check prices and bidding activity across multiple facilities via the webcast stream, simply

[6] To limit the possibility that these results are confounded by changes in geographic price dispersion associated with the webcast channel (discussed below), we reran the analysis using only the observations from the first six months of 2003, when webcast transactions were under 1% of total volume. Those results are similar to those we report.

[7] This gap exists for several reasons. One is to provide time for vehicles to be cleaned, reconditioned, and otherwise prepared for auction. Another is to allow time for similar vehicles to accumulate at a facility, as some sellers like to auction similar vehicles in sequence in an attempt to attract more buyers. Another is that vehicles don't always sell when first auctioned, which increases the time between distribution and sale.

by opening multiple browser windows. If he deems prices at one facility to be too high, then he can eschew bidding at that facility in favor of bidding at a different facility where prices are lower (if the auctions at the two facilities are conducted concurrently) or where he expects prices to be lower (if the auctions are not concurrent). This mitigates the need to forecast prices at each facility by providing buyers with real-time price information, and it increases the likelihood that buyers will shift their purchasing to facilities where prices are relatively low. Appendix A shows this via a simple proof.

To summarize, the webcast channel does not allow sellers to shift supply once prevailing prices become apparent (because that would require transporting vehicles), but it does allow buyers to shift demand. Another benefit of the webcast channel for buyers is that it eliminates the travel and opportunity costs of physical attendance at a facility. This reduction in participation costs should make buyers more likely to purchase at remote facilities, further helping them exploit geographic price differences. This is less of a benefit to sellers because their physical presence has never been strictly required at a facility; sellers can substitute for their physical attendance by providing a list of reserve prices to the auctioneer.

Given the benefits of the webcast channel to buyers, we posit that as the webcast channel diffuses, buyers will engage in more *price-driven remote purchases*. We define price-driven remote purchases as purchases in which the buyer purchases from a remote facility instead of a local facility because of lower prices. Price-driven remote purchases should reduce geographic price dispersion. The rationale is straightforward. Because prices are determined by auction, a buyer who shifts away from a local facility where prices are high to purchase at a remote facility where prices are low will increase the price at the latter facility by virtue of having outbid the other buyers. This will reduce the price dispersion between the two facilities. The buyer's shifting his demand away from the local facility might also *lower* the price there, further reducing the price dispersion, although this depends on the valuations of the other buyers at the local facility. Another way to think of this is that buyers who shift demand to low-price facilities reduce geographic price dispersion by bringing prices at those facilities closer to the mean. We summarize this via the following hypotheses.

Hypothesis 1 (H1). *The diffusion of the webcast channel is associated with an increase in price-driven remote purchases.*

Hypothesis 2 (H2). *An increase in price-driven remote purchases is associated with a decrease in geographic price dispersion.*

In the next section, we test H1 and H2 by examining how the webcast channel affected buyer behavior and the corresponding effect on geographic price dispersion. In §6, we consider the extent to which potential changes in seller behavior might also explain the reduction in geographic price dispersion.

5. Analysis of Buyer Behavior and Geographic Price Dispersion

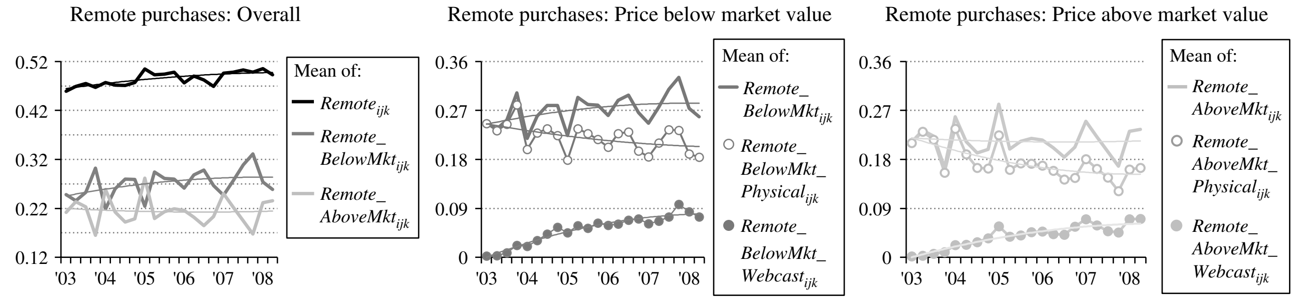
We begin this section by showing how trends in remote purchasing provide descriptive evidence for H1. We then present the results of different specifications of a discrete choice model, which provide more formal support for H1. Last, we show that growth in price-driven remote purchasing is associated with reduced geographic price dispersion (H2).

5.1. Temporal Trends in Price-Driven Remote Purchasing

As per H1, the diffusion of the webcast channel should lead to an increase in price-driven remote purchasing. To examine this, we set an indicator variable $Remote_{ijk} = 1$ for each *remote purchase*, defined as a purchase in which the buyer i was not "local" to the facility k where he purchased vehicle j. For our primary measure, we defined buyer i's "local" facility to be the one within 170 miles of his zip code at which he made the most physical purchases; we defined all other facilities as "remote" to buyer i.[8] We considered this intuitive, as this means that buyer i's local facility is a nearby facility that he visited frequently. We then decomposed $Remote_{ijk}$ into $Remote\_BelowMarket_{ijk}$ and $Remote\_AboveMarket_{ijk}$. We set $Remote\_BelowMarket_{ijk} = 1$ for transactions in which $Remote_{ijk} = 1$ and the vehicle's price was below its market value (see Table 1); $Remote\_BelowMarket_{ijk} = 0$ otherwise. $Remote\_AboveMarket_{ijk}$ is analogous. $Remote\_BelowMarket_{ijk} = 1$ represents a price-driven remote purchase because buyers are purchasing from remote facilities where prices are low. As discussed above, this should reduce price dispersion by bringing prices at

[8] As a secondary measure, we considered any facility within 170 miles of buyer i's zip code to be "local" to that buyer, with all other facilities "remote." We used 170 miles because buyers traveled an average of 170 miles when they made physical purchases. If there were no facilities within 170 miles of buyer i (which is true for 5.4% of buyers), we defined buyer i's local facility to be the closest facility. As tertiary and quaternary measures, we considered a buyer to be "local" to (a) the closest facility, irrespective of whether he visited that facility; and (b) the facility within 170 miles from which he purchased the most vehicles in the first quarter of 2003 (when webcast purchasing was in its infancy). For the latter measure, if there were no facilities within 170 miles of buyer i, we defined buyer i's local facility to be the closest facility. We also used 100, 150, and 200 as the mileage thresholds. Our results are robust to each measure.

Figure 3 **Remote Purchases (as a Proportion of Total Purchases) per Quarter**



Notes. Polynomial trend lines shown. See text for variable definitions.

low-priced facilities closer to the mean. We also decomposed *Remote_BelowMarket*$_{ijk}$ and *Remote_AboveMarket*$_{ijk}$ based on whether the purchases were made by buyers using the webcast or physical channels (using *_Webcast* and *_Physical* suffixes to label the decomposed variables).

Figure 3 plots the means of these variables by quarter; weekly plots are similar. The left panel shows that the percentage of remote purchases increased over time, with the upward trend driven by an increase in the percentage of price-driven remote purchases (*Remote_BelowMarket*$_{ijk}$). This shows that buyers became more likely over time to buy from remote facilities—*if* prices were low. Similar to the trend in price dispersion (see Figure 1), most of the increase in price-driven remote purchases occurred during the first half of the sample period, after which the slope of the increase flattened. We discuss this further in §5.4. The middle panel of Figure 3 shows that the upward trend in price-driven remote purchases (*Remote_BelowMarket*$_{ijk}$) was driven by an increase in price-driven remote purchases by buyers using the webcast channel (*Remote_BelowMarket_Webcast*$_{ijk}$) rather than by an increase by buyers using the physical channel. This suggests that the increase in price-driven remote purchases resulted from buyers' adoption of the webcast channel. The right panel of Figure 3 shows that remote purchases via the webcast channel also increased for vehicles above their market values. However, they increased 39% more rapidly (based on the slopes of linear regression trend lines) for vehicles below their market values; also, the mean of *Remote_BelowMarket_Webcast*$_{ijk}$ exceeded that of *Remote_AboveMarket_Webcast*$_{ijk}$ in 18 of the 22 quarters in the sample, including 12 of the last 13. Overall, the change in price-driven remote purchasing mirrored the changes in webcast purchasing and geographic price dispersion (see Figure 1). This suggests that buyers used the webcast channel to shift demand geographically to exploit price differences (H1) and that this reduced price dispersion (H2). We examine these correlations more formally below.

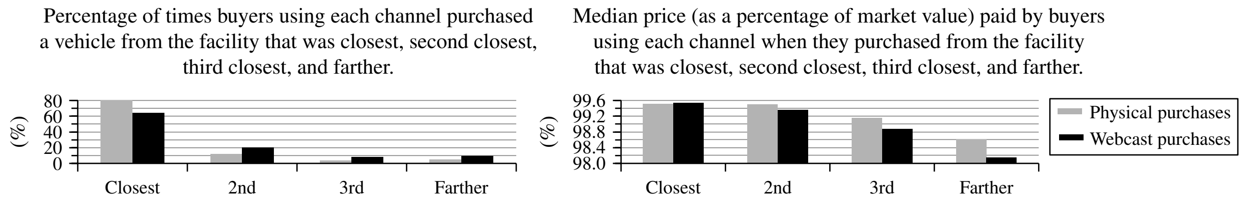
5.2. Discrete Choice Model of Buyer Behavior

We used discrete choice methods to test H1. The intuition behind this is as follows. Vehicles of year/model j are available at multiple facilities; buyers choose from among these facilities when deciding where to purchase. Buyers using the physical channel are likely to purchase from facilities close to them, given the costs of (a) traveling to a facility and (b) transporting vehicles back to their dealerships. Because buyers using the webcast channel bear the second cost but not the first, they will still prefer nearby facilities but will be more likely to purchase from remote facilities. Furthermore, because buyers using the webcast channel have the ability to shift their demand in response to real-time price information, they will be more likely to purchase from facilities where prices are relatively low (see §4.2). In a discrete choice framework, this will result in buyers who use the webcast channel appearing more sensitive to price but less sensitive to distance than buyers who use the physical channel. This would support H1 by providing evidence that webcast buyers engage in more price-driven remote purchasing than do physical buyers.

5.2.1. Descriptive Evidence. Before presenting the model, we present descriptive statistics to improve the transparency of our results and to limit the possibility that our results are artifacts of modeling or measurement assumptions. For each instance in which a buyer i purchased a vehicle(s) of year/model j at a facility k on day t, we defined the choice set K_{ijt} for that buyer as the set of facilities at which vehicles of year/model j were auctioned on day t.[9] In each instance, one of the K_{ijt} facilities was the closest to buyer i (based on his zip code), another was the second closest, etc. (Assuming K_{ijt} contains at least two alternatives.) The left panel of Figure 4 shows that buyers tended to purchase from

[9] To manage the scope of our analysis and to retain the focus on our research questions, we do not analyze *which* vehicles buyers decide to purchase. Instead, we assume that buyers know which vehicles they want to purchase, and we analyze where buyers choose to purchase them. This is consistent with research in this area (e.g., Chiou 2009, Jensen 2007).

Figure 4 Illustration of Differences in Where Buyers Purchased Vehicles (and How Much They Paid) When Using the Physical and Webcast Channels



Notes. Percentages shown in the left panel are conditional on K_{ijt} having at least four elements, i.e., that there were at least four facilities at which a vehicle(s) of year/model j was auctioned on day t. The pattern is the same without this condition, although there are (necessarily) fewer instances in which buyers chose the more remote facilities.

the closest facility. However, webcast buyers were more likely than physical buyers to purchase from a facility *other than* the closest. Buyers paid a smaller percentage of a vehicle's market value (based on *Valuation*) as they purchased from more distant facilities. However, this pattern is starker for buyers using the webcast channel (see right panel of Figure 4). This suggests that webcast buyers were more likely than physical buyers to shift demand to remote facilities, particularly when prices there were low, consistent with H1.

5.2.2. Model of Facility Choice Conditional on Channel. We modeled where buyer i chose to purchase vehicles of year/model j on day t from his choice set K_{ijt}, fitting one model for cases in which buyers used the physical channel and a separate model for cases in which buyers used the webcast channel. (Below, we endogenize the choice of channel by modeling the channel and facility choice jointly.) We modeled the utility of each facility in the choice set as follows:

$$\begin{aligned} U_{ijkt} = {} & \beta_{0,k} + \beta_1 \times PriceValRatio_{jkt} + \beta_2 \times Distance_{ik} \\ & + \beta_3 \times Supply_{jkt} + \beta_4 \times BuyerPropensity_{ijkt-} \\ & + \beta_5 \times Condition_{jkt} + \varepsilon_{ijkt}. \end{aligned} \tag{1}$$

In (1), $\beta_{0,k}$ are alternative-specific constants that capture the latent utility of purchasing at each facility k;[10] $Distance_{ik}$ is the number of miles between buyer i's zip code and facility k's zip code; and $Supply_{jkt}$ is the number of vehicles of year/model j auctioned at facility k on day t. As discussed in §§3 and 4, $Supply_{jkt}$ is determined prior to the buyer's choosing the facility at which to purchase, i.e., buyers and sellers do not arrive at facilities simultaneously. We include $Supply_{jkt}$ because buyers consider it when choosing a facility and to control for possible changes in how sellers distributed vehicles across facilities (which they might have done in response to implementation of the webcast channel; see §6). $BuyerPropensity_{ijkt-}$ is the proportion of all vehicles of year/model j that buyer i purchased in the 13 weeks (or fewer for observations in 2003) prior to day t that he purchased at facility k; $BuyerPropensity_{ijkt-}$ accounts for the possibility that buyers choose to buy vehicles where they have bought them in the past, perhaps out of habit or because of a relationship with the staff at facility k. We operationalized the other variables in the utility function slightly differently based on whether the facility was the chosen facility or a nonchosen facility in the choice set. For the chosen facilities, $PriceValRatio_{jkt}$ is the actual transaction price paid by the buyer divided by the vehicle's valuation, and $Condition_{jkt}$ is the condition grade assigned to the vehicle. For the nonchosen facilities, $PriceValRatio_{jkt}$ and $Condition_{jkt}$ represent values averaged across the vehicles of year/model j sold at facility k on day t. The rationale behind this distinction is that because we observe the actual price and condition grade of the vehicle that buyer i purchased at the chosen facility k, we use them in the model. We cannot do this for the nonchosen facilities, because we do not know which of the vehicles at a nonchosen facility the buyer would have purchased had he chosen that facility. Accordingly, we use the averages.[11] Buyers know $Distance_{ik}$ and

[10] Equivalently, $\beta_{0,k}$ are the coefficients for indicator variables that represent each facility k.

[11] Two comments are in order. First, a potential issue with this approach is that the values used to compute each instance of $PriceValRatio_{jkt}$ for the nonchosen alternatives may be widely dispersed. If that is the case, then $PriceValRatio_{jkt}$ could be an imprecise measure of the price of many of the vehicles at the nonchosen alternatives. To examine this, we computed the average standard deviation of the prices that form each instance of $PriceValRatio_{jkt}$, which is 0.02. As a robustness check, we restricted the analysis to choice cases in which the standard deviation of the prices that form $PriceValRatio_{jkt}$ for each nonchosen alternative was 0 ($n = 253{,}595$ choices for buyers using the physical channel and $n = 20{,}134$ choices for buyers using the webcast channel). This occurs when—for each nonchosen facility k on day t in the choice case—either only one vehicle of year/model j was sold or multiple vehicles were sold for the same $PriceValRatio_{jkt}$. This ensures that $PriceValRatio_{jkt}$ is a precise measure of the price at each nonchosen alternative in the choice set. These results are similar to the results from the full sample. In particular, β_1 is -0.373 (s.e. $= 0.049$) and -2.204 (s.e. $= 0.148$) for the physical and webcast channels, respectively, and β_2 is -0.007 (s.e. $= 0.000$) and -0.004 (s.e. $= 0.000$). Second, the specific price (and other variables) of nonchosen alternatives in choice models is often unobserved in the literature. Researchers often impute price for all nonchosen alternatives based on a hedonic regression or similar means (e.g., see Bucklin et al. 2008, p. 480; Chiou 2009, p. 292). We believe that our estimates of the nonchosen prices—which *are*

their own $BuyerPropensity_{ijkt-}$, and they can calculate $Supply_{jkt}$ and $Condition_{jkt}$ from the "presale" list of vehicles posted in advance on the intermediary's website. Thus, each of the variables in the model is known to the buyer when he chooses a facility, except for $PriceValRatio_{jkt}$, although $PriceValRatio_{jkt}$ is known to some extent by buyers using the webcast channel; see §4.2.[12] Table 2 provides descriptive statistics for these and other variables described later in this section.

We estimated the model (using a conditional logit specification) separately for the cases in which buyers purchased via the physical channel and in which they purchased via the webcast channel. Table 3 presents the results.

We tested H1 by comparing β_1 and β_2 across the channels. If buyers tended to purchase at facilities where prices were lower than at the other facilities in their choice sets, then β_1 will be negative. Similarly, if buyers tended to purchase at nearby facilities in their choice sets, then β_2 will be negative. As shown in Table 3, β_1 is more negative and β_2 is less negative for buyers using the webcast channel than for buyers using the physical channel. In other words, webcast buyers are more sensitive to price but less sensitive to distance. The 95% confidence intervals for β_1 and β_2 do not overlap across the two channels, and we also estimate the coefficients for both channels simultaneously in §5.2.3 to show their difference. To examine the economic significance of the differences across the channels, we used the model estimates for each channel to simulate how an increase in price at a buyer's "local" facility affects his propensity to purchase from a remote facility. We defined a buyer's "local" facility (a) using the primary measure from §5.1 and (b) as the facility in the

actual prices, averaged across one or more vehicles—are similar in precision (and potentially more precise) than many typically used in the literature.

[12] As with many models, ours is an approximation of the behavior of interest and should not be interpreted as a structural model of the buyers' behavior. Because of the auction setting, a structural model in our case would require modeling the equilibrium bid strategies of each buyer i for each vehicle of year/model j on each day t at each facility k, including how these strategies differ based on which channel buyer i is using and how many vehicles have already been (and are yet to be) auctioned. Rather than modeling the bid strategies, we assume that we observe the outcomes that result from buyers playing their optimal strategies. This greatly simplifies the analysis and keeps the focus of the paper on how use of the webcast channel affects geographic price dispersion. However, we recognize that modeling equilibrium bid strategies and the associated dynamics would be a major contribution to the structural auctions literature, in the vein of Jofre-Bonet and Pesendorfer (2003). Also, because $PriceValRatio_{jkt}$ is unknown to buyers using the physical channel when they choose a facility, the model is not structural. Despite this, the model is a useful statistical method for analyzing whether webcast buyers were more successful than physical buyers at purchasing from remote facilities where prices were relatively low, which is the goal of this aspect of the analysis.

Table 2 Variables Used in the Models of Buyers' Facility and Channel Choices (See §§5.2.2 and 5.2.3)

Variable	Mean (std. dev.)
$PriceValRatio_{jkt}$[a]	0.99 (0.11)[b]
$Distance_{ik}$	388 (339)
$Supply_{jkt}$	6.19 (11.33)
$BuyerPropensity_{ijkt-}$	0.04 (0.19)
$Condition_{jkt}$[c]	3.22 (0.66)
$NearbyFacilities_i$	3.95 (2.79)
$Webcast\_Selection_{kt}$	0.84 (0.36)
$Webcast\_Propensity_{it}$	0.09 (0.23)

Notes. Includes both chosen and nonchosen facilities in the choice sets. See text for variable definitions.

[a] $PriceValRatio_{jkt}$ is unobserved when none of the vehicles of year/model j auctioned at facility k on day t were sold. This occurs just over 10% of the time. We imputed the missing values as the average value of $PriceValRatio_{jkt}$ for the facilities at which vehicles of year/model j were sold on day t.

[b] This is the standard deviation of $PriceValRatio_{jkt}$ across all facilities in all choice sets. It should not be confused with the average standard deviation of the prices that comprise each instance of $PriceValRatio_{jkt}$, which is 0.02 (see Footnote 11).

[c] The condition grade of vehicles (*Condition*) is not recorded for approximately one-third of the vehicles. We imputed *Condition* for those vehicles based on (a) the average condition grade for vehicles of the same year/model j sold at the same facility k over the prior 21 days, (b) the vehicle's mileage, and (c) the vehicle's age. Specifically, we regressed *Condition* (when observed) on these three variables and used the resulting coefficients to impute *Condition* when not observed.

Table 3 Results of Model of Buyers' Choices of Facility Conditional on Channel

	Conditional on buyers using the physical channel	Conditional on buyers using the webcast channel[a]
	Coefficient	Coefficient
β_1: $PriceValRatio_{jkt}$	−0.398 (0.022)***	−1.767 (0.057)***
β_2: $Distance_{ik}$	−0.007 (0.000)***	−0.005 (0.000)***
β_3: $Supply_{jkt}$	0.056 (0.000)***	0.105 (0.001)***
β_4: $BuyerPropensity_{ijkt-}$	3.352 (0.011)***	2.618 (0.022)***
β_5: $Condition_{jkt}$	0.069 (0.003)***	0.264 (0.008)***
$\beta_{0,k}$: Facility constants	Included	Included
n (number of choices)[b]	1,272,190	170,696
Log likelihood	−595,158	−107,969

Notes. Models estimated using a conditional logit specification. Robust standard errors in parentheses.

[a] $PriceValRatio_{jkt}$, $Supply_{jkt}$, and $Condition_{jkt}$ for the nonchosen facilities in each choice set in this analysis were constructed using all vehicles of year/model j auctioned at facility k on day t. Because only vehicles in webcast-enabled lanes were available to buyers using the webcast channel, we also constructed these variables using only those vehicles of year/model j auctioned at facility k on day t in webcast-enabled lanes. If no vehicles of year/model j were auctioned in webcast-enabled lanes at facility k on day t, then we removed facility k from the choice set. These results are virtually identical to those we report.

[b] Differs from the total number of purchases made because (a) instances in which buyer i used a channel to purchase more than one vehicle of year/model j from facility k on day t were modeled as a single choice (see Appendix C) and (b) instances in which buyer i purchased a vehicle(s) of year/model j that was available at only one facility k on day t could not be included.

***$p < 0.01$.

Table 4 Simulated Percentage of Remote Purchases by Definition of "Remote," Simulated Scenario, and the Channel the Buyers Used

	% PurchasedRemotely: Primary measure		*% PurchasedRemotely*: Secondary measure	
	Baseline scenario	Alternative scenario[a]	Baseline	Alternative[a]
Buyers using the physical channel (%)	3.89	5.67	0.04	3.90
Buyers using the webcast channel (%)	5.00	18.03	0.01	21.97

Notes. The primary measure of *% PurchasedRemotely* is based on the primary definition of "remote" from §5.1. For the secondary measure, we defined "remote" as any facility in the choice set *other than* the one closest to the buyer.
[a]Same as the baseline scenario except $PriceValRatio_{jkt}$ is set one standard deviation above the mean for the buyer's local facility. See text for details.

choice set closest to the buyer. (For the first definition, we restricted the simulations to the choice cases in which the buyer's local facility was an alternative.) For each channel, we simulated two scenarios. In the baseline scenario, we set all variables (except $Distance_{ik}$) at their means. The alternative scenario is the same except we set $PriceValRatio_{jkt}$ one standard deviation above the mean for the buyer's local facility. As shown in Table 4, buyers using the webcast channel were three to six times more likely than buyers using the physical channel to purchase from a remote facility after the simulated price increase. This indicates that webcast buyers were more likely than physical buyers to purchase from remote facilities, particularly when prices at those facilities were relatively low. In other words, webcast buyers engaged in more price-driven remote purchasing (H1). Because more buyers used the webcast channel over time (see Figure 1), this helps explain the upward trend in price-driven remote purchases (see Figure 3).[13]

5.2.3. Model of Channel and Facility Choice. As an extension to the model, we endogenized the buyer's choice of channel. To model the joint channel/facility choice, we identified the set of facilities K_{ijt} at which vehicles of year/model j were auctioned on day t (as above). We assumed that buyer i could use either the physical or webcast channel to purchase from any of the K_{ijt} facilities, and so we included $2 \times K_{ijt}$ alternatives in each choice set. For example, if vehicles of year/model j are auctioned in Atlanta and Orlando on day t, then the choice set consists of two physical alternatives and two webcast alternatives: Atlanta—physical; Orlando—physical; Atlanta—webcast; and Orlando—webcast. We specified the utility of each facility/channel combination as $U_{ijckt} = \omega_{ickt} + \psi_{ijkt} + \varphi_{ijkt} + \varepsilon_{ijckt}$, where c denotes the channel, other subscripts are as above, and

$$\begin{aligned}\omega_{ickt} = {} & Webcast_c \times [\alpha_0 + \alpha_1 \times NearbyFacilities_i \\ & + \alpha_2 \times Webcast\_Selection_{kt} \\ & + \alpha_3 \times Webcast\_Propensity_{it-}], \end{aligned} \tag{2}$$

$$\begin{aligned}\psi_{ijkt} = {} & \beta_{0,k} + \beta_1 \times PriceValRatio_{jkt} + \beta_2 \times Distance_{ik} \\ & + \beta_3 \times Supply_{jkt} + \beta_4 \times BuyerPropensity_{ijkt-} \\ & + \beta_5 \times Condition_{jkt}, \end{aligned} \tag{3}$$

$$\begin{aligned}\varphi_{ijkt} = {} & Webcast_c \times [\gamma_{0,k} + \gamma_1 \times PriceValRatio_{jkt} \\ & + \gamma_2 \times Distance_{ik} + \gamma_3 \times Supply_{jkt} + \gamma_4 \\ & \times BuyerPropensity_{ijkt-} + \gamma_5 \times Condition_{jkt}]. \end{aligned} \tag{4}$$

$Webcast_c$ is an indicator variable set to 1 for the webcast alternatives in the choice set and 0 otherwise. The ω_{ickt} term represents the utility buyer i receives from purchasing via the webcast channel relative to the physical channel. Following prior literature about choice between physical and electronic channels (e.g., Forman et al. 2009), this utility depends on a constant (α_0), how accessible the physical market facilities are to buyer i ($NearbyFacilities_i$ is the number of facilities within 170 miles of buyer i), the selection of vehicles available via webcast ($Webcast\_Selection_{kt}$ is the proportion of vehicles at facility k on day t that were available via webcast), and how much buyer i has recently purchased via webcast ($Webcast\_Propensity_{it-}$ is the proportion of buyer i's 10 (or fewer if necessary) purchases prior to day t made via the webcast channel). $Webcast\_Selection_{kt}$ and $Webcast\_Propensity_{it-}$ capture the dynamics of the diffusion of the webcast channel; the former evolves as the webcast technology became more widely implemented at the facilities, whereas the latter evolves as buyers gained experience with the webcast channel. The ψ_{ijkt} term represents the utility of purchasing at facility k via the physical channel; it is the same as (1). The φ_{ijkt} term allows the utility of purchasing at facility k to shift if the buyer is using the webcast channel ($\varphi_{ijkt} = Webcast_c \times \psi_{ijkt}$ with different coefficients).[14]

[13] One might worry about a potential selection issue in which buyers who were already engaging in price-driven remote purchasing via the physical channel simply shifted their purchasing to the webcast channel. This is unlikely for four reasons. First, the webcast channel enables price-driven remote purchasing in a way that the physical channel does not, as described in §4.2. Second, if the result was caused by this selection issue, then price-driven remote purchasing would appear flat throughout the sample period as opposed to increasing with webcast channel use (see Figures 1 and 3). Third, the matching estimation reported in Appendix B provides evidence that the webcast channel increased price-driven remote purchasing as opposed to the two being simultaneously influenced by a third factor such as selection. Fourth, we limit potential selection issues by endogenizing the choice of channel in §5.2.3 and implementing the robustness check described in Footnote 16.

[14] One way to think of the model (but not the only way) is as a nested logit model in which the "upper" model represents a buyer's

Table 5 Results of Model of Buyers' Choices of Channel and Facility

	1: Conditional logit coefficient	2: Nested logit coefficient	3: Mixed logit coefficient
α_0: $Webcast_c$ *constant*	−8.192 (0.184)***	−12.010 (0.426)***	−12.768[a] (0.330)***
α_1: $Webcast_c \times NearbyFacilities_i$	−0.005 (0.006)	−0.016 (0.010)*	−0.031 (0.010)***
α_2: $Webcast_c \times Webcast\_Selection_{kt}$	4.552 (0.158)***	6.255 (0.253)***	6.019 (0.216)***
α_3: $Webcast_c \times Webcast\_Propensity_{it-}$	5.033 (0.024)***	8.137 (0.307)***	9.723 (0.231)***
β_1: $PriceValRatio_{jkt}$	−0.040 (0.057)	−0.042 (0.058)	−0.082 (0.065)
β_2: $Distance_{ik}$	−0.006 (0.000)***	−0.006 (0.000)***	−0.007 (0.000)***
β_3: $Supply_{jkt}$	0.065 (0.001)***	0.064 (0.001)***	0.073 (0.001)***
β_4: $BuyerPropensity_{jkt-}$	2.677 (0.028)***	2.666 (0.029)***	3.008 (0.037)***
β_5: $Condition_{jkt}$	0.081 (0.009)***	0.076 (0.009)***	0.084 (0.011)***
γ_1: $Webcast_c \times PriceValRatio_{jkt}$	−0.605 (0.083)***	−0.844 (0.123)***	−0.874 (0.127)***
γ_2: $Webcast_c \times Distance_{ik}$	0.002 (0.000)***	0.002 (0.000)***	0.003 (0.000)***
γ_3: $Webcast_c \times Supply_{jkt}$	0.011 (0.000)***	0.025 (0.002)***	0.021 (0.001)***
γ_4: $Webcast_c \times BuyerPropensity_{jkt-}$	−0.134 (0.019)***	0.026 (0.049)	−0.256 (0.032)***
γ_5: $Webcast_c \times Condition_{jkt}$	0.176 (0.014)***	0.238 (0.020)***	0.232 (0.021)***
$\beta_{0,k}$: $Facility(k)$	Included	Included	Included
$\gamma_{0,k}$: $Webcast_c \times Facility(k)$	Included	Included	Included
Inclusive value (physical nest)	n/a	0.682 (0.019)***	n/a
Inclusive value (electronic nest)	n/a	0.619 (0.023)***	n/a
n (number of choices)[b]	339,082	339,082	339,082
Log likelihood	−114,143	−114,063	−112,948

Notes. Each column shows the results from a different specification. Standard errors in parentheses.
[a]Modeled as a normally distributed random coefficient: Estimated standard deviation equals 2.92 (standard error equals 0.09).
[b]The sample size is smaller than that in Table 3 because this analysis is based on a geographic subset of the data, as discussed in Appendix C.
$^{*}p < 0.10$; $^{***}p < 0.01$.

Table 5 shows the results using different specifications. We used a conditional logit specification as a baseline. A limitation of this specification is that it has the independence from irrelevant alternatives ("IIA") property, which follows from the assumption that each of the error terms in the choice set are independent and identically distributed. IIA may be violated in our context, as buyers' unobservable channel preferences may create correlation in the errors for the physical alternatives and for the webcast alternatives. We addressed this by using a nested logit specification in which we allowed the errors for the two types of alternatives to be correlated by placing the physical and webcast alternatives into separate nests. In addition to unobserved heterogeneity in *channel* preferences that create correlation in the errors, buyers might also have unobserved heterogeneity in *facility* preferences that create correlation in the errors for the two alternatives (physical and webcast) for a given facility. To address this, we used a mixed logit specification in which we modeled α_0 and $\beta_{0,k}$ as normally distributed random coefficients.[15]

The constant term for $Webcast_c$ (α_0) is negative and significant, which is consistent with the majority of purchases being conducted in the physical channel. The positive and significant coefficients for $Webcast_c \times Webcast\_Selection_{it}(\alpha_2)$ and $Webcast_c \times Webcast\_Propensity_{it-}$ (α_3) indicate that the utility of the webcast channel to a buyer increased as more vehicles were made available via webcast and as he gained experience using it. The coefficient for $PriceValRatio_{jkt}$ (β_1) is negative but insignificant, whereas the coefficient for $Webcast_c \times PriceValRatio_{jkt}$ (γ_1) is negative and significant. This indicates that buyers were more likely to purchase from facilities where prices were relatively low when using the webcast channel than when using the physical channel. The coefficient for $Distance_{ik}$ (β_2) is negative and significant, whereas the coefficient for $Webcast_c \times Distance_{ik}$ (γ_2) is positive and significant. (Note that the "combined" coefficient $\beta_2 + \gamma_2 = -0.004$ ($p < 0.01$) represents the effect of $Distance_{ik}$ for webcast buyers.) This indicates that buyers preferred nearby facilities when using the webcast channel, but less so than when using the physical channel.[16]

choice of channel and the "lower" model represents a buyer's choice of facility. In this interpretation, α_0 represents the alternative-specific constant for the webcast channel in the upper model.

[15] The random coefficients capture unobserved heterogeneity in buyer preferences for channels and facilities, but more importantly, they yield an overlapping nest structure that allows the error terms to be correlated across the alternatives for each channel and each facility. This is because including the random coefficients is equivalent to specifying normally distributed error components for each channel and facility (as shown in Train 2009, p. 140).

[16] Buyers who used the physical channel may differ from buyers who used the webcast channel, and this heterogeneity might explain part of our results. To examine this, we refitted the models presented in §5.2.2 using only the choice cases for buyers who used both

To summarize, buyers were more likely to purchase from a remote facility where prices were relatively low when using the webcast channel than when using the physical channel (H1). Buyers became more likely to use the webcast channel as more vehicles became available via webcast and as their recent experience with the webcast channel grew. This explains why the prevalence of price-driven remote purchasing increased (see Figure 3) with the diffusion of the webcast channel (see Figure 1).

In Appendix B, we use an alternative strategy (matching estimation) to test H1 and find corroborating results.

5.3. Relationship of Price-Driven Remote Purchases to Geographic Price Dispersion

We next examine whether increasing price-driven remote purchasing was associated with decreasing geographic price dispersion (H2). To examine (and estimate the size of) this relationship, we used a fixed effects panel regression model (i.e., the "within" estimator) that uses the temporal changes in remote purchasing and geographic price dispersion to identify the coefficients. Equation (5) shows our base specification.

$$\begin{aligned} &\textit{GeoPriceDisp\_CoefVar}_{jt} \text{ or } \textit{GeoPriceDisp\_StDev}_{jt} \\ &\quad = \beta_0 + \beta_1 \times \textit{RemotePurchases}_{jt} + \beta_2 \times \textit{LocalPurchases}_{jt} + \tau_t \\ &\qquad + c_j + \varepsilon_{jt}, \quad \text{where } \tau_t \text{ are week fixed effects and} \\ &\qquad c_j \text{ are fixed effects for vehicle age/model } j.^{17} \end{aligned} \tag{5}$$

RemotePurchases$_{jt}$ (*LocalPurchases*$_{jt}$) is the number of purchases for vehicles of age/model j in week t in which the buyer was *not* local (was local) to the facility where he purchased the vehicle. We used our primary measure of whether a facility was "local" to the buyer; we obtain similar results using our secondary measures (see §5.1). In some regressions, we decomposed *RemotePurchases*$_{jt}$ into *RemotePurchases_BelowMarket*$_{jt}$ and *RemotePurchases_AboveMarket*$_{jt}$ based on whether vehicles were purchased for below or above market value, as in §5.1. This decomposition allows us to examine the differential effect of price-driven remote purchases. We did the same for *LocalPurchases*$_{jt}$. These and other variables used in the regression are described in Table 6. Regression results appear in Table 7. Including the week fixed effects means that our results take into account the dramatic increase in transport costs (see Figure 1 and §5.4) as well as unobserved temporal trends such as changes in how sellers distributed vehicles over time, which we explore more directly in §6.

Table 7 shows that an increase in remote purchases is significantly associated with a decrease in geographic price dispersion (β_1), and that this relationship is specific to *price-driven* remote purchases (β_{1b}). (β_{1b} and β_{1a} are statistically different at $p < 0.01$ in all versions of the model.) This provides support for H2. A one standard deviation increase in *RemotePurchases_BelowMarket*$_{jt}$ is associated with an approximately 8% decrease in geographic price dispersion. Both β_{1b} and β_{2b} are negative and significant (and statistically different at $p < 0.01$). This indicates that both remote and local "below-market" purchases are associated with decreased price dispersion, with β_{1b} having the larger magnitude and marginal effect. (Marginal effects are straightforward given the linearity of the model.) Related, both β_{1a} and β_{2a} are positive and significant, likely because "above-market" purchases increase geographic price dispersion by sending above-average prices farther from the mean.

5.4. Summary of Why Increases in Price-Driven Remote Purchasing Led to Lower Geographic Price Dispersion

Our analysis shows that the webcast channel helped buyers shift their demand across facilities to exploit price differences and that this led to reduced geographic price dispersion. The growth in use of the webcast channel over time—and the corresponding growth in price-driven remote purchases—mirrored the reduction in price dispersion. Figure 5 illustrates why increasing levels of price-driven remote purchasing led to decreasing levels of geographic price dispersion. Essentially, the decrease in geographic price dispersion was small when only a few buyers engaged in price-driven remote purchasing. As more buyers engaged in price-driven remote purchasing, geographic price dispersion decreased further.

Of course, there is a limit to how "remotely" buyers will purchase because they must pay to transport vehicles to their locations. This explains why buyers' sensitivity to distance is negative when using the webcast channel, albeit less negative than when using the physical channel. The cost of transporting vehicles also means that geographic price dispersion will never be zero. Instead, there will be a baseline level, consistent with

channels for their recent purchases, specifically, those with a value of *Webcast_Propensity*$_{it-}$ between 0.3 and 0.7. These results are consistent with those from the main analysis, which makes this alternative explanation unlikely. Specifically, β_1 is -0.238 (s.e. $= 0.096$) and -1.934 (s.e. $= 0.108$) for the physical and webcast channels, respectively, and β_2 is -0.007 (s.e. $= 0.000$) and -0.005 (s.e. $= 0.000$).

[17] We estimated the model using both age/model (e.g., one-year-old Chevy Malibu; here, age is measured in integers) and year/model (e.g., 2002 Chevy Malibu) as the panel variable. We report the results for age/model; results for year/model are similar. A drawback to using year/model as the panel variable is that observations of older model year vehicles (e.g., 2001 and before) become sparse in the latter part of the sample period. As a result, the *number* of remote purchases declines over time for these vehicles, even though the *proportion* of remote purchases typically increases. Using age/model as the panel variable solves this issue by providing better balance within the panel.

Table 6 Variables Used to Analyze the Relationship Between Remote and Local Purchasing and Geographic Price Dispersion

Variable	Description	Mean (std. dev.)
$GeoPriceDisp\_CoefVar_{jt}$	Geographic price dispersion of vehicles of age/model j in week t. Computed by (a) calculating the mean price of vehicles of age/model j at each facility k in week t, referred to as *facility means*; and (b) calculating the coefficient of variation of the facility means for age/model j in week t.	0.11 (0.11)
$GeoPriceDisp\_StDev_{jt}$	Same as $GeoPriceDisp\_CoefVar_{jt}$, except with the standard deviation as the dispersion measure instead of the coefficient of variation.	1,578 (1,471)
$GeoPriceMean_{jt}$	Mean of the facility means (see above) for age/model j in week t.	16,976 (8,323)
$RemotePurchases_{jt}$	Number of purchases for vehicles of age/model j in week t in which the buyer was *not* local to the facility where he purchased the vehicle.	6.26 (15.52)
$LocalPurchases_{jt}$	Number of purchases for vehicles of age/model j in week t in which the buyer was local to the facility where he purchased the vehicle.	6.71 (19.63)
$RemotePurchases\_BelowMarket_{jt}$	Number of remote purchases for vehicles of age/model j in week t in which the vehicle sold for less than its average market value.	3.36 (8.65)
$RemotePurchases\_AboveMarket_{jt}$	Number of remote purchases for vehicles of age/model j in week t in which the vehicle sold for at or above its average market value.	2.90 (7.32)
$LocalPurchases\_BelowMarket_{jt}$	Same as $RemotePurchases\_BelowMarket_{jt}$, except for local purchases.	2.90 (9.94)
$LocalPurchases\_AboveMarket_{jt}$	Same as $RemotePurchases\_AboveMarket_{jt}$, except for local purchases.	3.82 (10.10)

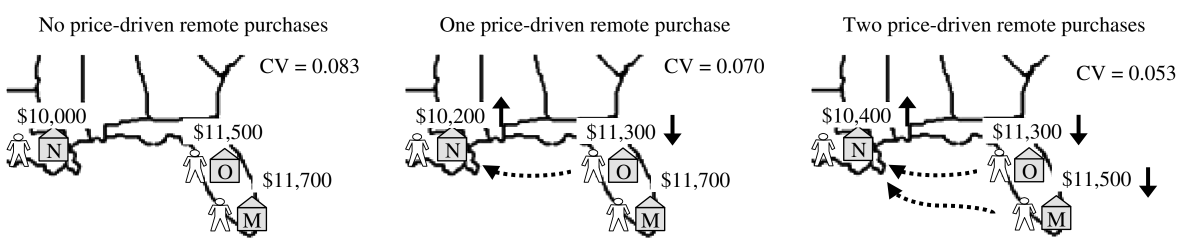
Table 7 Relationship Between Remote and Local Purchasing and Geographic Price Dispersion

	Dep. var.: $GeoPriceDisp\_CoefVar_{jt}$		Dep. var.: $GeoPriceDisp\_StDev_{jt}$	
β_1: $RemotePurchases_{jt}$	−0.0002 (0.0000)***	—	−3.30 (0.39)***	—
β_{1b}: $RemotePurchases\_BelowMarket_{jt}$	—	−0.0011 (0.0001)***	—	−13.49 (1.12)***
β_{1a}: $RemotePurchases\_AboveMarket_{jt}$	—	0.0006 (0.0001)***	—	6.34 (0.93)***
β_2: $LocalPurchases_{jt}$	−0.0000 (0.0000)	—	0.20 (0.27)	—
β_{2b}: $LocalPurchases\_BelowMarket_{jt}$	—	−0.0005 (0.0000)***	—	−5.32 (1.30)***
β_{2a}: $LocalPurchases\_AboveMarket_{jt}$	—	0.0007 (0.0000)***	—	8.03 (1.04)***
β_3: $GeoPriceMean_{jt}$	—	—	−0.02 (0.01)***	−0.02 (0.01)***
β_0: $Intercept$	0.0854 (0.0118)***	0.0848 (0.0118)***	1,538 (160)***	1,519 (160)***
Week fixed effects	Yes	Yes	Yes	Yes
Age/model fixed effects	Yes	Yes	Yes	Yes
R^2 (w/fixed effects)	0.28	0.28	0.25	0.25
n	161,132	161,132	161,132	161,132

Note. Robust standard errors in parentheses.
$^{***}p < 0.01$.

the law of one price, which allows prices to vary up to the cost of transport. As discussed in §4, this baseline level likely increased over the sample period due to the dramatic increase in fuel costs (see Figure 1). To explore this, we reran the panel regression reported above after replacing the week fixed effects with the weekly price of diesel fuel per gallon (denoted $FuelPrice_t$; mean = 2.42, s.d. = 0.75; from EIA 2014). The coefficient for $FuelPrice_t$ is 0.0057 (s.e. = 0.0004) and 106.6 (s.e. = 5.31) with $GeoPriceDisp\_CoefVar_{jt}$ and $GeoPriceDisp\_StDev_{jt}$, respectively, as the dependent variable. In both cases, a one standard deviation increase is associated with a 5% increase in price dispersion. The coefficient and standard error for $RemotePurchases\_BelowMarket_{jt}$ are

Figure 5 Illustration of Why Geographic Price Dispersion Decreases with Increasing Levels of Price-Driven Remote Purchasing



Notes. "CV" is the coefficient of variation of prices across facilities (N, New Orleans; O, Orlando; M, Miami) in each scenario. In the first scenario, all buyers (represented by the person icons) purchase at their local facilities. In the second scenario, there is one price-driven remote purchase—represented by the dashed arrow—which lowers the geographic price dispersion. In the third scenario, there are two price-driven remote purchases, which lowers the geographic price dispersion by a greater degree.

virtually identical to those shown in Table 7; all other results are similar as well. This means that the upward trend in fuel prices was *increasing* geographic price dispersion at the same time the upward trend in price-driven remote purchases was *decreasing* it. These dueling effects help explain why the downward trend in price dispersion (see Figure 1) flattened after the first part of the sample period, when fuel prices began their climb to historic highs. Also, as transport costs increase, the price differences across facilities must be larger for price-driven remote purchasing to be rational. This helps explain the flattening of the upward trend in price-driven remote purchases (see Figure 3) over the latter part of the sample period. Overall, it appears that the market's adherence to the law of one price increased over time, even as the geographic price dispersion trend flattened.

6. Analysis of Seller Behavior

The webcast channel helps buyers shift demand between facilities in response to real-time price information, which reduces geographic price dispersion. The webcast channel provides no analogous benefit to sellers. However, shifts in seller behavior during the sample period might account for at least some of the reduction in geographic price dispersion, which we explore here.

The relevant seller behavior for this analysis is how sellers choose the facilities at which to sell vehicles because this determines the supply at each facility, which affects prices.[18] We examined this via a discrete choice model. Each auctioned vehicle represents a choice made by the seller to offer that vehicle at a given facility. Although sellers can distribute vehicles to any facility in the country, none of the sellers in the data use all facilities. We assume that each seller s has a set of facilities K_s at which she offers vehicles; we recover these sets for each seller based on where we observe them to offer vehicles and use them as the choice sets. We modeled the utility to seller s of choosing facility k for a vehicle year/model j that is auctioned on day t as

$$\begin{aligned}
U_{jkst} = {} & \beta_{0,k} + \beta_1 \times Supply_{kt-} + \beta_2 \times Supply^2_{kt-} \\
& + \beta_3 \times Supply_PctWebcastAvailable_{kt-} \\
& + \beta_4 \times Purchases_LocalBuyers_{kt-} \\
& + \beta_5 \times Purchases_LocalBuyers_PctWebcast_{kt-} \\
& + \beta_6 \times PriceValRatio_{jkt-} + \beta_7 \times SellerPropensity_{jkst-} \\
& + \beta_8 \times SellerPropensity_{jkst-} \times Prevalence_{jst} \\
& + \beta_9 \times SellerPropensity_{jkst-} \times GeoPriceStDev_{jt-} \\
& + \beta_{10} \times SupplyYearModel_{jkt-} \\
& + \beta_{11} \times SupplyYearModel^2_{jkt-} \\
& + \beta_{12} \times PctSold_{jkt-} + \varepsilon_{jkst}.
\end{aligned}$$

For consistency with prior research, we included all variables shown to affect seller distribution behavior in this market (Overby and Clarke 2012). We extended the model from Overby and Clarke (2012) by adding the $Supply\_PctWebcastAvailable_{kt-}$ and $Purchases\_LocalBuyers\_PctWebcast_{kt-}$ variables to examine the effect of the implementation of the webcast channel, and we also explored sellers' sensitivity to recent prices in greater detail. We used lagged variables because contemporaneous variables are unknown to sellers when they make distribution decisions, which they typically do weeks before the vehicles are auctioned on day t (see §4.2). As such, lagged variables approximate the information that sellers have when making decisions. We varied the lag length for robustness; the reported results are based on a 13-week lag.[19] We describe each variable in Table 8.

Another variable that influences sellers' utility is the distance between a vehicle g's location prior to its entering the market and facility k, which we label $Distance_{gk}$. Recall that there are two types of sellers in the market: commercial sellers and dealer sellers. For commercial sellers, we have no data on where their vehicles are located prior to being brought into the market; $Distance_{gk}$ is unobserved for these sellers. However, we can infer $Distance_{gk}$ for dealer sellers by assuming that they store vehicles at their dealerships prior to bringing them into the market. This is a reasonable assumption given that dealer sellers want these vehicles available to potential retail customers and that dealerships are designed to store vehicles. Because dealers' zip codes are recorded in the data when they purchase a vehicle, we computed $Distance_{gk}$ as the distance between the zip code of the dealer who offered vehicle g and facility k. We included $Distance_{gk}$ in the utility function for this subset of the data, which helped us examine whether the inclusion/exclusion of $Distance_{gk}$ affects our other estimates. We estimated the model using a conditional logit specification. Results appear in Table 9.

[18] Another seller behavior is whether they accept or reject the high bid for vehicles. However, it is unlikely that changes in this behavior could explain the reduction in geographic price dispersion. For this to be a valid explanation, sellers would have to set consistent reserve prices across the country. We cannot rule this out because seller reserve prices are unobserved, but we believe it to be unlikely given that it would require (a) a significant level of coordination/collusion among the thousands of sellers in the market and (b) this coordination/collusion to have become more pronounced over time to explain the downward trend in geographic price dispersion.

[19] Results using other lag lengths are similar, although the magnitude and marginal effect of the $PriceValRatio_{jkt-}$ coefficient gets closer to zero with shorter lags. This is consistent with Overby and Clarke (2012), who used a three-week lag and found a positive but insignificant effect of price (using a different sample).

Table 8 Variables Used in the Model of Sellers' Distribution Choices

Variable	Description	Mean (std. dev.)
$Supply_{kt-}$	Number of vehicles auctioned at facility k in the 13 weeks prior to week t. Also represents a measure of the size of the facility.	3,323 (2,693)
$Supply\_PctWebcastAvailable_{kt-}$	Proportion of vehicles auctioned at facility k auctioned in webcast-enabled lanes in the 13 weeks prior to week t.	0.77 (0.34)
$Purchases\_LocalBuyers_{kt-}$	Number of purchases (at all facilities) by buyers local to facility k in the 13 weeks prior to week t. We defined "local" as in §5.1.	1,961 (1,328)
$Purchases\_Local\ Buyers\_PctWebcast_{kt-}$	Proportion of purchases (at all facilities) by buyers local to facility k made via the webcast channel in the 13 weeks prior to week t.	0.11 (0.09)
$PriceValRatio_{jkt-}$	Average price to valuation ratio for vehicles of year/model j sold at facility k in the 13 weeks prior to week t.[a]	0.98 (0.13)[b]
$SellerPropensity_{jkst-}$	Proportion of all vehicles of year/model j that seller s offered in the 13 weeks (or fewer for observations in 2003) prior to week t that she offered at facility k.	0.07 (0.17)
$Prevalence_{jst}$	Proportion of all vehicles that seller s offered in week t that were of year/model j.	0.18 (0.23)
$GeoPriceDisp\_StDev_{jt-}$	Geographic price dispersion of vehicles of year/model j in the three weeks prior to week t. Calculated by (a) computing the mean price of vehicles of year/model j at each facility they were sold in the three weeks prior to week t (labeled the *facility means*) and (b) computing the standard deviation of the facility means from (a).[c, d]	1,562 (760)
$SupplyYearModel_{jkt-}$	Number of vehicles of year/model j auctioned at facility k in the 13 weeks prior to week t. Set to 0 if no vehicles of year/model j were auctioned.	48.5 (99.3)
$PctSold_{jkt-}$	Proportion of vehicles of year/model j auctioned at facility k in the 13 weeks prior to week t that were sold. Set to 0 if no vehicles of year/model j were auctioned.	0.51 (0.38)[e]
$Distance_{gk}$	Distance between the zip codes of the seller who offered vehicle g and facility k. Available for dealer sellers only.	326 (573)

Note. Descriptive statistics include both chosen and nonchosen facilities in the choice sets.

[a] $PriceValRatio_{jkt-}$ is null if no vehicles of year/model j were sold at facility k in the 13 weeks prior to week t. In these cases, we replaced the nulls with a version of $PriceValRatio_{jkt}$ calculated using only the vehicle model, the vehicle year and make, or the vehicle make (in that order). We set remaining nulls to the average value of $PriceValRatio_{jkt}$ for the vehicles of year/model j that were sold in the 13 weeks prior to week t at *any* facility.

[b] This is the standard deviation for $PriceValRatio_{jkt-}$ across all facilities in all choice sets. It should not be confused with the average standard deviation of the prices that form each instance of $PriceValRatio_{jkt-}$, which is 0.02.

[c] We used a three-week lag because a longer lag might obscure price dispersion by averaging across too long of a period.

[d] $GeoPriceDisp\_StDev_{jt-}$ is null when vehicles of year/model j were sold at one or fewer facilities over the lagged period. In these cases, we calculated $GeoPriceDisp\_StDev_{jt-}$ by extending the lagged period until it covered two facilities where vehicles of year/model j were sold. We dropped the cases for which this did not work (0.6%).

[e] The mean is lower than the percentage of all vehicles in the data that sold because $PctSold_{jkt-} = 0$ if no vehicles of year/model j were auctioned at facility k during the lagged period.

We examined the $Supply\_PctWebcastAvailable_{kt-}$ (β_3) and $Purchases\_LocalBuyers\_PctWebcast_{kt-}$ (β_5) coefficients to consider whether the diffusion of the webcast channel affected the geographic distribution of supply in the market (and thereby potentially the price dispersion). First, sellers favored facilities at which a high degree of vehicles were available via the webcast channel (β_3 is positive and significant), likely because those facilities could attract additional buyers. Second, sellers *disfavored* facilities at which the buyers local to that facility purchased a high percentage of vehicles via webcast (β_5 is negative and significant). This is likely because sellers preferred facilities where the local buyers were "captive" and unlikely to use the webcast channel to shift their demand to another facility. $Supply\_PctWebcastAvailable_{kt-}$ and $Purchases\_LocalBuyers\_PctWebcast_{kt-}$ are positively correlated ($\rho = 0.46$), likely because buyers local to heavily "wired" facilities received substantial exposure to the webcast channel, thereby hastening their adoption of it. We used the model estimates for the full sample to simulate whether the opposing effects of these two variables would offset. We simulated the percentage change in the number of times facility k was chosen when $Supply\_PctWebcastAvailable_{kt-}$ and $Purchases\_LocalBuyers\_PctWebcast_{kt-}$ were one standard deviation below their means vs. when they were at their means. We did this for each facility; the average percentage change was minimal: −0.2% (std. dev. 0.6%). Thus, these variables—which reflect the diffusion of the webcast channel—appear to have offsetting effects on the geographic distribution of supply.

The coefficient for $PriceValRatio_{jkt-}$ (β_6) is positive and significant. At first glance, this may suggest that sellers added supply to facilities where prices were already above market value, which would lower prices at those facilities and thereby account for some of the reduction in geographic price dispersion. However, as shown in §4.1, prices fluctuate substantially from week to week, and lagged price premiums (discounts) are a poor predictor of whether a vehicle will sell for a premium (discount) at the time it is auctioned. Thus, a positive β_6 means only that sellers considered lagged prices when distributing vehicles; it does not mean that sellers

Table 9 Results of the Model of Sellers' Distribution Choices

	Full sample	Subsample for which we inferred $Distance_{gk}$	
	Coefficient	Coefficient	Coefficient
β_1: $Supply_{jkt-}$ [a]	0.232 (0.027)***	−0.289 (0.055)***	−0.305 (0.055)***
β_2: $Supply^2_{jkt-}$	−0.335 (0.016)***	−0.036 (0.030)	−0.103 (0.030)
β_3: $Supply\_PctWebcastAvailable_{kt-}$	0.110 (0.007)***	0.031 (0.014)**	0.054 (0.014)***
β_4: $Purchases\_LocalBuyers_{kt-}$ [a]	0.139 (0.029)***	0.102 (0.060)	0.276 (0.060)***
β_5: $Purchases\_LocalBuyers\_PctWebcast_{kt-}$	−0.663 (0.026)***	−0.751 (0.050)***	−0.759 (0.050)***
β_6: $PriceValRatio_{jkt-}$	0.381 (0.010)***	0.056 (0.011)***	0.060 (0.011)***
β_7: $SellerPropensity_{jkst-}$	4.138 (0.008)***	4.672 (0.015)***	4.570 (0.015)***
β_8: $SellerPropensity_{jkst-} \times Prevalence_{jst}$	−0.909 (0.012)***	−1.763 (0.017)***	−1.669 (0.017)***
β_9: $SellerPropensity_{jkst-} \times GeoPriceDisp\_StDev_{jt-}$ [a]	−0.125 (0.004)***	−0.192 (0.006)***	−0.193 (0.006)***
β_{10}: $SupplyYearModel_{jkt-}$ [a]	0.461 (0.003)***	0.510 (0.006)***	0.520 (0.006)***
β_{11}: $SupplyYearModel^2_{jkt-}$	−0.045 (0.000)***	−0.067 (0.001)***	−0.071 (0.001)***
β_{12}: $PctSold_{jkt-}$	0.377 (0.003)***	0.242 (0.005)***	0.241 (0.005)***
β_{13}: $Distance_{gk}$	—	—	−0.004 (0.000)***
$\beta_{0,k}$: *Facility* (k)	Included	Included	Included
n (number of choices)[b]	2,717,217	859,977	859,977
Log likelihood	−3,935,528	−846,551	−839,549

Note. Standard errors in parentheses.
[a]To ensure that coefficients were of similar magnitude for reporting purposes, we divided several variables by multiples of 10 prior to estimating the model. We divided $Supply_{jkt-}$ by 10,000, $Purchases\_LocalBuyers_{kt-}$ by 10,000, $GeoPriceDisp\_StDev_{jt-}$ by 1,000, and $SupplyYearModel_{jkt-}$ by 100.
[b]Differs from the total number of vehicles auctioned because (a) sellers who always distributed vehicles to the same facility could not be included and (b) instances in which we could not calculate $GeoPriceDisp\_StDev_{jt-}$ were dropped (see Table 8).
$^{**}p < 0.05$; $^{***}p < 0.01$

successfully exploited geographic price differences by sending vehicles to facilities where prices were high at the time they were auctioned. Even if we consider a positive β_6 to be evidence that sellers were successful at distributing additional supply to high-priced facilities, this is unlikely to explain the reduction in price dispersion we observe for two reasons. First, sellers would have had to become more sensitive to prices over time (as captured by an increase in the marginal effect of β_6) to explain the reduction in price dispersion. In other words, sellers' distribution behavior would have to *change* during the sample period to explain the *change* in price dispersion. We reestimated the model by quarter and found that β_6 (and the marginal effect of a one standard deviation increase, estimated via simulation as above) does not increase over time. Instead, the quarter-by-quarter trend for β_6 is essentially flat. Second, the practical significance of β_6 is low. Using simulation (as above), a one standard deviation increase in $PriceValRatio_{jkt-}$ yields a 2.8% increase in the number of vehicles distributed to a facility, on average. The number of vehicles of year/model j auctioned at facility k in week t is four or less 91% of the time and the mean is 2.3 (s.d. = 4.8).[20] Given that vehicles can only be distributed in whole numbers, a 2.8% increase for a year/model j that would otherwise have four vehicles at facility k will result in one additional vehicle being auctioned at facility k every nine weeks, on average (i.e., $1/(4 * 0.028) \approx 9$). In other words, instead of the number of vehicles of year/model j at facility k over a nine-week span being 4-4-4-4-4-4-4-4-4, it would be 4-4-4-4-4-4-4-4-5, and *only* if $PriceValRatio_{jkt-}$ stayed one standard deviation above its mean for a nine-week period, which is unlikely (see §4). Overall, these results provide little evidence that sellers were successful at increasing (decreasing) the supply at facilities where prices were high (low), which might have explained some of the reduction in geographic price dispersion.

The most powerful explanatory variable (based on log-likelihood ratio tests) is $SellerPropensity_{jkst-}$, which is consistent with prior research (Overby and Clarke 2012). This indicates that the best predictor of where a seller will distribute a vehicle is where she has distributed vehicles of that year/model in the past. The high explanatory power of this variable is likely because sellers must make distribution decisions on a recurring basis (as often as daily or weekly), and they may simplify their decision making by relying on habit or by leveraging existing relationships with the management at a facility. Some sellers may also have agreements to supply vehicles to certain facilities. This indicates that there are limits to how strategic sellers are when distributing vehicles, at least in terms of trying to exploit geographic price differences, which is consistent with our findings immediately above.

The $Distance_{gk}$ coefficient shown in the third column of Table 9 is negative and significant, indicating that

[20] These statistics are smaller than those for $Supply_{jkt}$ listed in Table 2. That is because the statistics in Table 2 are averaged across facilities in the buyers' choice sets. Facilities with large values for $Supply_{jkt}$ appear in more choice sets and are therefore included more times in the statistics reported in Table 2.

sellers prefer nearby facilities, ceteris paribus. Also, the similarity of the coefficients in the second and third columns of Table 9 show that the inclusion/exclusion of $Distance_{gk}$ has minimal effect on the other coefficients. Although we cannot be sure that this holds for the full sample, it suggests that the inability to observe $Distance_{gk}$ for the full sample may not seriously bias our coefficients.

7. Conclusion

In many geographically distributed markets, traders lack information about prices across market locations, and it can be costly to conduct transactions at remote locations. These frictions may prevent goods from being allocated efficiently, leading to high geographic price dispersion and violation of the "law of one price." Because electronic commerce improves price visibility and lowers transaction costs, increasing levels of electronic trading may improve efficiency and reduce geographic price dispersion. We study this by analyzing how the diffusion of an electronic "webcast" channel affected geographic price dispersion in the wholesale used vehicle market. The webcast channel allows buyers to monitor market conditions at different facilities without having to travel, which helps them shift demand across market facilities in real time as prices at those facilities are revealed. The shifting of demand across facilities—which became more prevalent as the webcast channel diffused—reduced geographic price dispersion and caused the market to more closely reflect the law of one price.

7.1. Limitations

First, although we find little evidence that changes in seller behavior contributed to the observed reduction in geographic price dispersion, we cannot completely rule this out, particularly given the five-plus year span of our study. A potential supply-side effect might reduce the precision of our estimates of how much changes in buyers' price-driven remote purchasing affected geographic price dispersion (see §5.3). But it would not affect our overall conclusion that buyers used the webcast channel to exploit geographic price differences, which led to reduced geographic price dispersion. Second, we measured geographic price dispersion by comparing prices of vehicles of the same year/model across locations. This ensures that the vehicles are highly similar, and we further limited vehicle heterogeneity by restricting the sample to vehicles with between 15,000 and 21,000 miles. However, no two vehicles in the sample are identical because of differences in color, options, etc. If the unobserved heterogeneity within vehicle year/models lessened over time, then this might explain some of the reduction in geographic price dispersion. We cannot rule this out.

7.2. Contributions

Markets can improve the efficiency of exchange, but they operate best when information flows freely and transaction costs are low. Our study contributes to the literature on how new technologies such as electronic channels can improve market function by improving access to information and lowering transaction costs. We extend prior research in this area in several ways. First, we examine and document the microlevel behavioral mechanism by which electronic trading affects price dispersion, which is critical for continued empirical research in this stream (Baye et al. 2006). We use trend analysis, discrete choice methods, fixed effects panel regression, and matching estimation to show that buyers used the electronic "webcast" channel to shift demand from local facilities where prices were high to remote facilities where prices were low, causing a reduction in geographic price dispersion. Second, we show that the location of buyers and products influenced how buyers used the physical and electronic channels and the facilities at which they purchased. Location has often been overlooked in the literature on electronic commerce and price dispersion, despite the fact that it plays a critical role in the trade of many products such as automobiles, food, and raw materials. Third, we examine at a microlevel how the diffusion of the electronic channel affected not only where buyers chose to buy but also where sellers chose to sell. This extends prior work that has drawn inferences directly from price data (Aker 2010) or examined the behavior of only one party (Jensen 2007, Overby and Clarke 2012).

Acknowledgments

The authors thank the Networks, Electronic Commerce, and Telecommunications Institute [Grant 11-30] and the Georgia Institute of Technology for financial support. This research has benefited from valuable comments from Lorin Hitt (department editor), the associate editor, the referees, and seminar participants at the Georgia Institute of Technology, the University of Minnesota, Purdue University, Michigan State University, the 2009 International Conference on Information Systems, the 2009 Statistical Challenges in E-Commerce Research Symposium, and the 2012 Networks, Electronic Commerce, and Telecommunications Institute Conference.

Appendix A. Simple Proof of Why Webcast Buyers Are More Likely Than Physical Buyers to Purchase from Facilities Where Prices Are Relatively Low

This proof extends the argument in §4.2 about why webcast buyers are more likely than physical buyers to purchase from facilities where prices are relatively low.

Setup: Assume that vehicles of year/model j (e.g., 2007 Honda Accord) are auctioned on day t at two facilities, A and B, with the auctions occurring first at facility A. We denote the prices at each facility (denoted p_A and p_B) as either below market value (L) or above market value (H). Let the probability that prices at facility A are below market value be $P(p_A = L) = \psi$ such that $P(p_A = H) = 1 - \psi$. Analogously, $P(p_B = L) = \zeta$ and $P(p_B = H) = 1 - \zeta$. The following table lists the possible scenarios for the prices at the two facilities, the probability of each, and the probability that $p_A < p_B$ in each scenario. By definition, $P(p_A < p_B) = 1$ in scenario 2 and $P(p_A < p_B) = 0$ in scenario 3. In scenarios 1 and 4, we assume that $P(p_A < p_B) = 0.5$.

Scenario	Price at A (p_A)	Price at B (p_B)	Probability of scenario	Probability that $p_A < p_B$
1	$p_A = L$	$p_B = L$	$\psi \times \zeta$	0.5
2	$p_A = L$	$p_B = H$	$\psi \times (1-\zeta)$	1
3	$p_A = H$	$p_B = L$	$(1-\psi) \times \zeta$	0
4	$p_A = H$	$p_B = H$	$(1-\psi) \times (1-\zeta)$	0.5

Assume that a buyer purchases a vehicle of year/model j on day t at either facility A or B. The buyer can use either the physical or webcast channel. As described in §4.2, if the buyer uses the physical channel, then he must commit to facility A or B before the auctions begin, i.e., before he can assess whether prices at facility A are below or above market value. By contrast, if the buyer uses the webcast channel, then he does not have to commit. Instead, he can observe whether prices at facility A are above or below market value and use that information to choose between A and B.[21] Assume that the buyer uses the following decision rule when using the webcast channel. If prices at facility A are below market value, i.e., if $p_A = L$, then he will purchase at A; if prices at facility A are above market, i.e., if $p_A = H$, then he will defer and purchase at B.

Proposition. *The buyer is more likely (or at least as likely) to purchase at the lower-priced facility when using the webcast channel than when using the physical channel.*

Proof. We consider the two exhaustive cases: the buyer either purchases at facility A (case A) or he purchases at facility B (case B). We begin with case A. When the buyer is using the physical channel, the probability that he purchases at the lower-priced facility, i.e., $P(p_A < p_B)$, is

$$\begin{aligned}
&P(p_A < p_B \mid physical)\\
&\quad = P(scenario1) \times P(p_A < p_B) + P(scenario2) \times P(p_A < p_B)\\
&\qquad + P(scenario3) \times P(p_A < p_B) + P(scenario4) \times P(p_A < p_B)\\
&\quad = 0.5(\psi - \zeta + 1).
\end{aligned}$$

When the buyer is using the webcast channel, he only purchases at facility A (i.e., case A) if $p_A = L$. Thus, $P(p_A = L) = \psi = 1$. Because $\psi = 1$, we can write

$$\begin{aligned}
&P(p_A < p_B \mid webcast)\\
&\quad = P(scenario1) \times P(p_A < p_B) + P(scenario2) \times P(p_A < p_B)\\
&\qquad + P(scenario3) \times P(p_A < p_B) + P(scenario4) \times P(p_A < p_B)\\
&\quad = 0.5(2 - \zeta).
\end{aligned}$$

[21] The webcast buyer can determine whether prices for vehicles at facility A are above or below market value by referencing the vehicles' published *Valuations* (see Table 1).

Case B is analogous, yielding $P(p_B < p_A \mid physical) = 0.5(\zeta - \psi + 1)$ and $P(p_B < p_A webcast) = 0.5(\zeta + 1)$.

The proof is completed by showing—in both cases A and B—that the probability that the buyer purchases at the lower-priced facility is greater (or the same) when using the webcast channel than the physical channel, i.e., that

Case A:	Case B:
$P(p_A < p_B \mid webcast)$ $\geq P(p_A < p_B \mid physical)$.	$P(p_B < p_A \mid webcast)$ $\geq P(p_B < p_A \mid physical)$.
$0.5(2 - \zeta) \geq 0.5(\psi - \zeta + 1)$.	$0.5(\zeta + 1) \geq 0.5(\zeta - \psi + 1)$.
$\psi \leq 1$.	$\psi \geq 0$.

Because ψ is a probability, the condition holds for both cases. □

Appendix B. Testing H1 via Matching Estimation

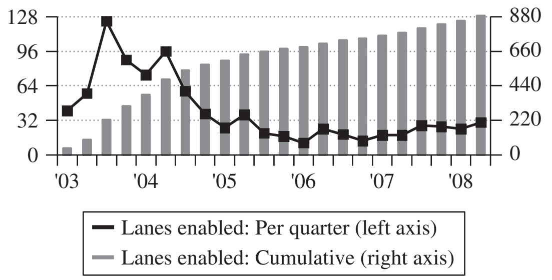
In the main text, we presented the results of one econometric strategy for examining the effect of the webcast channel on price-driven remote purchasing (H1): a discrete choice model. We corroborated these results via a second econometric strategy: a matching estimator. These two strategies exploit different types of variation in the data. The discrete choice model exploits differences in purchasing behavior across the physical and webcast channels, and the matching estimator exploits the phased nature of the implementation of the webcast channel. The two strategies complement one another; we present them both to show that our findings are robust to assumptions required by either.

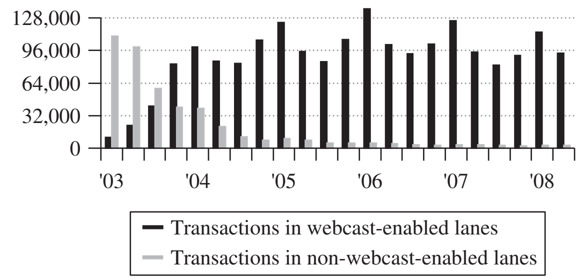
As discussed in §3.1.2, the webcast technology was implemented in different lanes at different times. We received a second data source from the intermediary that allowed us to estimate when the webcast channel was implemented in each lane at each facility. These data include all vehicles auctioned in all lanes at all facilities between January 2003 and June 2008. For each lane/facility combination (i.e., lane 1 in Houston, lane 2 in Houston, etc.) we recorded the date of the first webcast purchase and used that as the webcast implementation date.[22] The phased implementation of the webcast channel means that during the implementation, each facility had some lanes that were webcast-enabled and some that were not. The left panel of Figure B.1 shows that the bulk of the implementation occurred in 2003 and 2004. Consistent with this, the right panel of Figure B.1 shows that from 2005 onward, most of the transactions in the sample occurred in webcast-enabled lanes, although most transactions continued to be conducted by buyers using the physical channel (see Figure 1); note that webcast-enabled lanes are accessible by buyers using either the physical or webcast channels.

Because of the phased implementation, we observe many instances in which highly similar vehicles were sold at the same facility on the same day, some of which were available via webcast and some of which were not. We considered each vehicle sold in a webcast-enabled lane to be a potential "treated" vehicle and each vehicle sold in a non-webcast-enabled lane to be a potential "control" vehicle. We used a

[22] Because there could be a lag between webcast implementation and the date of the first webcast purchase for a lane, we reran our analysis after adjusting the webcast implementation date by subtracting one week, three weeks, and six weeks. This does not affect our results.

Figure B.1 (Left Panel) Lanes Enabled with the Webcast Technology per Quarter; (Right Panel) Transactions That Occurred in Webcast- and Non-Webcast-Enabled Lanes per Quarter





matching procedure to generate a set of treated and control vehicles that are essentially equivalent except that the treated vehicles were sold in webcast-enabled lanes. If matched appropriately, the control vehicles serve as counterfactuals for the treated vehicles, such that differences in transaction outcomes can be attributed to the "treatment effect" of webcast enablement (Imbens 2004). The matching procedure allows us to test whether treated vehicles were more likely to be purchased by a remote buyer than were control vehicles, particularly when they sold for below market value. In this way, we can test whether the webcast channel increased price-driven remote purchasing beyond what would have otherwise occurred.

The Matching Algorithm: We matched treated vehicles to control vehicles using exact matching and coarsened exact matching (Iacus et al. 2011a, b). Each treated vehicle could only be matched to a control vehicle with the same year/model (e.g., 2003 Ford Ranger) sold at the same facility (e.g., Boston). We coarsened *Valuation* (see Table 1) into $1,000 bins and only matched vehicles within the same bin. Combined with the exact matching on vehicle year/model, this ensures that matched vehicles are highly similar.[23] We coarsened *AuctionDate* by week and only matched vehicles sold in the same week. The matching algorithm produces a cell for each combination of attributes used in the matching. We retained cells with *at least* one treated and one control vehicle (some retained cells have more); this is required to estimate the treatment effect and is standard for matching estimators (Imbens 2004). The remaining observations comprise the matched sample, which consists of 91,848 matched vehicles in 13,854 cells.

Model Specification and Results: We fitted the following regression model on the matched sample to test H1: $Remote_{ijk} = \beta_0 + \beta_1 \times WebcastEnabled_{jk} + \varepsilon_{ijk}$, with $Remote_{ijk}$ defined as above. We used the primary measure of $Remote_{ijk}$ for our focal analysis (see §5.1); results are robust to alternative measures. $WebcastEnabled_{jk}$ is an indicator variable set to 1 if the vehicle was offered in a webcast-enabled lane (i.e., was treated) and 0 otherwise. The model allows us to test whether webcast treatment increased the probability that a vehicle was purchased by a remote buyer. We also replaced $Remote_{ijk}$ with $Remote\_BelowMarket_{ijk}$ and $Remote\_AboveMarket_{ijk}$ to examine whether the webcast treatment effect differed based on whether the vehicle sold for below or above market value. We fitted the regressions using weighted least squares to account for the (sometimes) unequal number of treated and control vehicles within each cell.[24] Linear regression results for the different dependent variables are shown in Table B.1, both for all matches and for only those matches from 2003 and 2004 (given the few control vehicles available after 2004). Logistic regression results are similar.

Table B.1 shows that webcast treatment is associated with a significant increase in the probability that a vehicle is purchased by a remote buyer, but only when the vehicle sells for below market value. Webcast treatment is associated with a *decrease* in the probability of a remote buyer purchasing a vehicle for *above* market value, although this effect is insignificant for the full matched sample. In unreported analysis, we find that these results hold when the sample is restricted to the "best" matches, i.e., those for which the treated and control vehicles in each cell were sold on the same day and were within $100 of *Valuation* (on average). The increase in probability of a remote buyer purchasing the vehicle for below market value is between 8.8% and 21.8%, depending on which matched cells we include in the analysis. The *decrease* in probability of a remote buyer purchasing the vehicle for *above* market value is between 2.3% and 10.4%, although this decrease is often insignificant.[25] These results indicate that webcast enablement fostered price-driven remote purchasing, such that increasing levels of enablement (see Figure B.1) led to increasing levels of price-driven remote purchasing (see Figure 3). This supports H1.[26]

[23] A t-test shows that *Valuation* does not statistically differ between treated ($\mu_{\text{valuation}} = 13{,}837$) and control ($\mu_{\text{valuation}} = 13{,}836$) vehicles ($p = 0.31$). We did not match on *Mileage* because (a) *Mileage* is already controlled for by our sampling strategy and (b) *Mileage* is a major determinant of *Valuation*. A t-test shows that *Mileage* does not statistically differ between treated ($\mu_{\text{mileage}} = 18{,}109$) and control ($\mu_{\text{mileage}} = 18{,}111$) vehicles ($p = 0.82$).

[24] Following Iacus et al. (2011a), we assigned each treated vehicle a weight of 1 and each control vehicle a weight of $(m_c/m_t)/(m_t^S/m_c^S)$. The variables m_c and m_c^S are the number of control vehicles in the matched sample and in matched cell S, respectively, and m_t and m_t^S are analogous for the treated vehicles. In addition to the Iacus et al. (2011a) paper, readers interested in why the weights capture the treatment effect correctly are referred to King (2012).

[25] In all cases, we verified that the $WebcastEnabled_{jk}$ coefficients (β_1) for the $Remote\_BelowMarket_{ijk}$ and $Remote\_AboveMarket_{ijk}$ regressions were statistically different by estimating both regressions simultaneously using seemingly unrelated regression and rejecting the equality of the coefficients via a χ^2 test ($p < 0.01$ in all cases).

[26] Webcast enablement increased remote purchasing of vehicles that sell for below market value, but any associated geographic price

Table B.1 Effect of Webcast Enablement on Whether a Vehicle Was Purchased by a Remote Buyer, Including Whether It Sold for Below or Above Its Market Value

	Dependent variable		
	$Remote_{ijk}$	$Remote\_BelowMarket_{ijk}$	$Remote\_AboveMarket_{ijk}$
Using all matches			
β_0: *Constant*	0.4275 (0.0070)***	0.2052 (0.0054)***	0.2223 (0.0059)***
β_1: *WebcastEnabled* ("treated")	0.0287 (0.0073)***	0.0363 (0.0057)***	−0.0076 (0.0061)
n	91,848	91,848	91,848
Regression F-statistic	15.43***	40.85***	1.56
R^2	0.01	0.01	0.01
Using matches from 2003 and 2004 only			
β_0: *Constant*	0.4375 (0.0083)***	0.2068 (0.0064)***	0.2306 (0.0071)***
β_1: *WebcastEnabled* ("treated")	0.0236 (0.0087)***	0.0419 (0.0067)***	−0.0183 (0.0073)**
n	68,731	68,731	68,731
Regression F-statistic	7.43***	39.45***	6.22**
R^2	0.01	0.01	0.01

Note. Robust standard errors shown in parentheses.
$^{**}p < 0.05$; $^{***}p < 0.01$.

Model Assumptions and Robustness: A condition for valid matching estimation is unconfoundedness (also known as selection on observables), which is that treatment is independent of the outcome conditional on the matching variables. Stated differently, unconfoundedness assumes that there are no unobserved variables that systematically affect both whether the unit of observation receives the treatment and the outcome(s) of interest (see Imbens 2004). We believe that our data allows matches that are suitably unconfounded, as discussed below.

Whether a vehicle received the treatment—i.e., whether it was auctioned in a webcast-enabled lane—was a choice made by the managers at each facility when they determined how to assign vehicles to lanes. Ideally, this choice would have been made randomly, but as with most economic settings involving real-world data, it was not. Fortunately, we have knowledge of how this choice was made that we use to consider whether our matches are likely to be confounded.[27] As a general rule, management chose to "treat" vehicles that were likely to be attractive to remotely located buyers. This attractiveness depended on (a) the availability of a vehicle's year/model at other facilities and (b) a vehicle's quality uncertainty. First, managers chose to treat vehicles whose year/model was not widely available at other facilities. These vehicles are inherently likely to attract remote buyers, and treating these vehicles would make it easier for remote buyers to purchase them. This does not confound our analysis, however, because year/model is observed and controlled for directly by our exact matching. Second, managers chose to treat vehicles of low quality uncertainty because those vehicles could be purchased without physical inspection by the buyer (Overby and Jap 2009). All vehicles in our sample—both treated and control—have similarly low quality uncertainty. This is because they are all "lightly" used (given their low mileage), such that significant quality issues are unlikely to have developed. Even if some vehicles have unobserved quality issues, managers are unlikely to be aware of this information because in-depth mechanical inspections are not standard practice. Even if management had perfect information, it's not clear why management would sort "lemons" into either the webcast-enabled or non-webcast-enabled lanes, as they have no incentive to engender adverse selection in any of their lanes. Thus, we are doubtful that unobserved quality distinguishes treated from control vehicles within our sample.

Of course, the unconfoundedness assumption is untestable, and we cannot rule out the possibility that unobserved variables inherently make vehicles chosen for treatment more attractive to remote buyers. This might confound a conclusion that webcast treatment increases the probability that a vehicle is purchased by a remote buyer. However, our conclusion is more nuanced than that. We find that webcast treatment increases the probability of a remote buyer when the vehicle's price is below market value but *not* when the price is above market value. This finding is consistent with our conclusion that buyers use the webcast channel to shift their demand geographically to exploit price differences. By contrast, it seems less likely for an omitted variable such as unobserved quality to have this type of nuanced effect.

In our main results, we restricted matches between treated and control vehicles to those of the same year, model, facility, and week and to those within the same $1,000 *Valuation* bin. This yields highly precise matches, but it also eliminates observations because treated (control) observations are dropped if they are not matched to a control (treated) observation (this is true of all matching estimators). To examine whether our matching results held over a wider range of the data, we rematched treated and control vehicles using less

dispersion reduction could be countervailed if webcast enablement also increased *local* purchasing of vehicles that sell for *above* market value. To examine this, we reran our regressions using the dependent variables $Local\_BelowMarket_{ijk}$ and $Local\_AboveMarket_{ijk}$, where $Local_{ijk}$ is coded opposite to $Remote_{ijk}$. The treatment effect (represented by β_1) is negative for both variables (and insignificant when using the matches from 2003–2004 only).

[27] This knowledge is based on interviews with management and the corresponding author's personal experience. The corresponding author consulted with the intermediary on the planning and initial implementation of the webcast system in 2002 and has interfaced with management on countless occasions since then.

Table B.2 Matching Estimation Results Based on Different Matching Criteria

Matched on					Dependent variable					
					$Remote_{ijk}$		$Remote\_BelowMarket_{ijk}$		$Remote\_AboveMarket_{ijk}$	
Time period	Model/ facility	Vehicle year	Valuation ($1,000 bins)	n	β_0: Constant	β_1: Webcast	β_0: Constant	β_1: Webcast	β_0: Constant	β_1: Webcast
Week	Yes	Yes	Yes	91,848	0.427***	0.029***	0.205***	0.036***	0.222***	−0.008
Month	Yes	Yes	Yes	225,175	0.418***	0.032***	0.196***	0.041***	0.222***	−0.008**
Quarter	Yes	Yes	Yes	367,309	0.420***	0.037***	0.200***	0.038***	0.220***	0.000
Week	Yes	Yes	No	130,468	0.435***	0.023***	0.213***	0.029***	0.222***	−0.005
Month	Yes	Yes	No	313,121	0.413***	0.042***	0.194***	0.045***	0.219***	−0.003
Quarter	Yes	Yes	No	518,741	0.410***	0.050***	0.201***	0.042***	0.208***	0.008***
Week	Yes	No	Yes	103,847	0.435***	0.023***	0.216***	0.026***	0.219***	−0.003
Month	Yes	No	Yes	255,308	0.418***	0.034***	0.201***	0.038***	0.217***	−0.004
Quarter	Yes	No	Yes	421,832	0.419***	0.040***	0.205***	0.036***	0.214***	0.003
Week	Yes	No	No	163,575	0.430***	0.028***	0.217***	0.027***	0.214***	0.001
Month	Yes	No	No	368,806	0.417***	0.040***	0.205***	0.039***	0.212***	0.002
Quarter	Yes	No	No	589,567	0.415***	0.046***	0.210***	0.037***	0.205***	0.009***

$^{**}p < 0.05$; $^{***}p < 0.01$.

precise criteria. First, we allowed matches between treated and control vehicles sold within the same month and quarter (see the "Time period" column of Table B.2). Second, we did not require vehicles to be matched on *VehicleYear* (see the "Vehicle year" column). Third, we allowed vehicles to be matched regardless of *Valuation*. Table B.2 shows the results from all permutations of these criteria. The pattern of results holds across the different matching criteria.

Appendix C. Additional Details About the Discrete Choice and Panel Regression Models

Discrete Choice Models: This section of the appendix contains robustness checks and some technical details about the discrete choice models reported in §§5.2.2 and 5.2.3.

First, for the models reported in §5.2.2, we limited the facilities in each choice set to those within 1,500 miles of the buyer's zip code. Less than 1% of purchases in the data were from facilities greater than 1,500 miles from the buyer. Excluding these (almost) never-chosen and likely never-considered facilities from the choice set prevents the model estimates from being perturbed by a large number of virtually never-chosen facilities in each set. Results are similar if we drop this restriction.

Second, for the instances in which buyer i purchased more than one vehicle of year/model j at facility k on day t, we averaged across the vehicles he purchased to construct $PriceValRatio_{jkt}$ and $Condition_{jkt}$ for the chosen facility. Our results are unaffected if we exclude those instances from the analysis.

Third, we reran the models by year to examine how the coefficients evolved over time; we found them to be stable. The $PriceValRatio_{jkt}$ coefficients for years 2003–2008 for the physical (webcast) channel were −0.34, −0.50, −0.52, −0.45, −0.34, −0.23 (−1.74, −1.62, −1.47, −1.92, −1.62, −1.84). The $Distance_{ik}$ coefficients for the physical (webcast) channel were −0.007 (−0.005) in all years. All coefficients for both variables were significant at $p < 0.05$. The mean and standard deviations for both variables were similar across all years.

Fourth, the high number of parameters (including the alternative-specific constants for each facility) in the joint channel/facility model discussed in §5.2.3 creates substantial dimensionality. This makes model estimation unstable and convergence difficult, particularly for specifications other than the conditional logit. To achieve an estimable model, we took a geographic subset of the sample comprised of only the purchases made by buyers local to the facilities in the western Uniited States, which consists of the facilities in Arizona, California, Colorado, Nevada, New Mexico, Oregon, Utah, and Washington. We used our primary definition of "local" (see §5.1) for this analysis. (Results also hold if we define each buyer's "local" facility to be the facility closest to him.) We also limited the choice sets to facilities within this region. This subsample is large enough geographically to allow us to examine whether buyers extended their geographic reach when purchasing via the webcast channel. Also, 97.3% of purchases by buyers local to facilities within this region are from facilities within this region, which allows us to consider it a microcosm of the entire market. The filtered choice data set consists of 339,082 choices and 18 facilities.

Fixed Effects Panel Regression Model: This section of the appendix contains robustness checks and some technical details about the fixed effects panel regression model reported in §5.3.

First, there is a risk of reverse causality in the panel regression model in the sense that low geographic price dispersion might lead to fewer price-driven remote purchases because there would be little incentive to shift demand across facilities. However, if this were the direction of the effect, then there would be a *positive* relationship between $GeoPriceDisp\_CoefVar_{jt}$ and $RemotePurchases\_BelowMarket_{jt}$, which would make it more difficult to recover the negative relationship shown in Table 7. Also, the analysis of buyer behavior provides a clear mechanism through which causality should flow from price-driven remote purchasing to price dispersion. Nevertheless, for robustness against this potential endogeneity concern, we instrumented $RemotePurchases\_BelowMarket_{jt}$ with its one-period lag. In the first-stage regression, $RemotePurchases\_BelowMarket_{jt-1}$ (i.e., the one-period lag) is positively correlated with $RemotePurchases\_BelowMarket_{jt}$ ($\beta = 0.13$, s.e. $= 0.00$) and the regression F-statistic is 1,983 ($p < 0.001$). After instrumentation, the β_{1b} coefficient becomes -0.0014 (s.e. $= 0.0003$) with $GeoPriceDisp\_CoefVar_{jt}$ as the dependent variable and -19.19 (s.e. $= 4.95$) with $GeoPriceDisp\_StDev_{jt}$ as the dependent variable.

Second, thinly traded vehicles have less temporal variation in the number of transactions per week than do thickly traded vehicles. To assess whether this limited variation might unduly influence the coefficients, we reran the regressions using only those vehicle age/models j that were traded at least n times per week (on average), with $n = 5$ and $n = 18$. Setting $n = 5$ ($n = 18$) restricts the sample to the vehicle age/models at or above the 70th (90th) percentile in terms of trading frequency. Results are similar to those we report.

References

Aker JC (2010) Information from markets near and far: Mobile phones and agricultural markets in Niger. *Amer. Econom. J.: Appl. Econom.* 2(3):46–59.

Baye MR, Morgan J, Scholten P (2006) Information, search, and price dispersion. Hendershott T, ed. *Handbook on Economics and Information Systems* (Elsevier, Amsterdam), 323–376.

Blum B, Goldfarb A (2006) Does the Internet defy the law of gravity? *J. Internat. Econom.* 70(2):384–405.

Brynjolfsson E, Smith MD (2000) Frictionless commerce? A comparison of Internet and conventional retailers. *Management Sci.* 46(4):563–585.

Bucklin RE, Siddarth S, Silva-Risso J (2008) Distribution intensity and new car choice. *J. Marketing Res.* 45(3):473–486.

Chellappa RK, Sin RG, Siddarth S (2011) Price-formats as a source of price dispersion: A study of online and offline prices in the domestic U.S. airline markets. *Inform. Systems Res.* 22(1):83–98.

Chiou L (2009) Empirical analysis of competition between Wal-Mart and other retail channels. *J. Econom. Management Strategy* 18(2):285–322.

Clemons EK, Hann IH, Hitt LM (2002) Price dispersion and differentiation in online travel: An empirical investigation. *Management Sci.* 48(4):534–549.

EIA (U.S. Energy Information Administration) (2014) Gasoline and diesel fuel update. Report, http://www.eia.gov/petroleum/gasdiesel/.

Forman C, Ghose A, Goldfarb A (2009) Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management Sci.* 55(1):47–57.

Hortaçsu A, Martínez-Jerez FA, Douglas J (2009) The geography of trade in online transactions: Evidence from eBay and MercadoLibre. *Amer. Econom. J.: Microeconomics* 1(1):53–74.

Iacus SM, King G, Porro G (2011a) Causal inference without balance checking: Coarsened exact matching. *Political Anal.* 20:1–24.

Iacus SM, King G, Porro G (2011b) Multivariate matching methods that are monotonic imbalance bounding. *J. Amer. Statist. Assoc.* 106(493):345–361.

Imbens G (2004) Nonparametric estimation of average treatment effects under exogeneity: A review. *Rev. Econom. Statist.* 86(1):4–29.

Jensen R (2007) The digital provide: Information (technology), market performance, and welfare in the south Indian fisheries sector. *Quart. J. Econom.* 122(3):879–924.

Jofre-Bonet M, Pesendorfer M (2003) Estimation of a dynamic auction game. *Econometrica* 71(5):1443–1489.

King G (2012) An explanation for CEM weights. Report, https://docs.google.com/document/d/1xQwyLt_6EXdNpA685LjmhjO20y5pZDZYwe2qeNoI5dE/edit?pli=1.

McMillan J (2002) *Reinventing the Bazaar: A Natural History of Markets* (Norton & Co, New York).

NAAA (National Auto Auction Association) (2013) NAAA 2013 annual review. Report, http://www.naaa.com/pdfs/2013AnnualReview.pdf.

NADA (National Automobile Dealers Association) (2012) NADA data 2012. Report, http://www.nada.org/NR/rdonlyres/C1C58F5A-BE0E-4E1A-9B56-1C3025B5B452/0/NADADATA2012Final.pdf.

Overby E, Clarke J (2012) A transaction-level analysis of spatial arbitrage: The role of habit, attention, and electronic trading. *Management Sci.* 58(2):394–412.

Overby E, Jap S (2009) Electronic and physical market channels: A multiyear investigation in a market for products of uncertain quality. *Management Sci.* 55(6):940–957.

Sheskin DJ (2004) *Handbook of Parametric and Nonparametric Statistical Procedures*, 3rd ed. (Chapman & Hall, Boca Raton, FL).

Stigler G (1961) The economics of information. *J. Political Econom.* 69(3):213–225.

Train KE (2009) *Discrete Choice Methods With Simulation*, 2nd ed. (Cambridge University Press, New York).