

# The Impact of Online Information on the Purchase of Certified Used Cars

Vandana Ramachandran  
Robert H. Smith School of Business  
University of Maryland, College Park, MD 20742.

Siva Viswanathan  
Robert H. Smith School of Business  
University of Maryland, College Park, MD 20742.

Sanjay Gosain  
Capital Group Inc.  
Irvine, CA.

## ABSTRACT

Secondary markets have adopted a number of quality signaling mechanisms such as certification to reduce information asymmetries between buyers and sellers in these markets. However the importance and value of these signals depends on the nature and extent of information asymmetries. With the growth of the Internet, consumers seeking to purchase used goods now have access to a plethora of information on various aspects of their purchase process. What then is the impact of such information on the salience and value of a traditional quality signal such as certification? We draw upon a unique and extensive dataset of consumers who obtained vehicle and transaction related information from online sources in their used vehicle purchase process to examine the impact of their information-acquisition on the choice of certification, as well as the price paid. We compare the outcomes of sales where consumers purchased certified used cars with sales of used-cars where there was no certification. Our findings highlight that product-related information substitutes, and price-related information complements, certification as indicated by their differential impacts on demand and price of certified used cars. We discuss the relevance of our findings for buyers and sellers and outline implications for online information providers as well.

**Keywords:** certification, online information, quality signals, uncertainty reduction, used cars

## Introduction

Most markets are characterized by significant information asymmetries between buyers and sellers, with sellers possessing superior information to buyers about the true quality of their offerings. As noted in the seminal paper by Akerlof (1970), when buyers lack adequate information to distinguish higher quality from lower quality products, sellers of low quality products have an incentive to misrepresent the quality of their offerings leading eventually to a market collapse. To avoid such a collapse, markets characterized by information asymmetries, particularly markets for used goods, have resorted to quality-signaling mechanisms such as certification that can help reduce friction in these markets. While such quality signaling mechanisms are valuable to both buyers and sellers in such markets, their value to consumers, and the competitive advantage they provide to sellers, depends crucially on the nature and extent of the information asymmetries in these markets. As noted by Folta and Janney (2004), signaling mechanisms are more valuable and result in higher returns when buyers face greater difficulty in acquiring information about underlying product attributes.

Over the recent years the growth of the Internet has dramatically increased the availability of information to prospective buyers across a number of markets. Consumers now, have access to a wide range of information regarding products, services, and alternatives, to make purchase decisions. Such information can potentially help bridge the information gap between buyers and sellers, and significantly reduce the information asymmetries in these markets. The abundance of information available through various online sources brings to fore questions regarding the salience and value of conventional quality signaling mechanisms such as certification. For instance, *are consumers more likely to purchase certification if they have more information? Does certification command a higher premium when consumers have more information? Do consumers who obtain certain types of information pay lower prices than others for their purchases?* The used-car market provides an excellent test-bed to examine these questions, and is one that provides the context for our study. Given the complexity of the offerings and the difficulty in determining quality, certification, in particular, plays a

very important role in reducing the frictions inherent in the market for used cars. While certification has traditionally been valuable to buyers as well as sellers of used cars, the growth of the Internet, and the emergence of auto-retailing websites in particular, has dramatically increased the amount of information available to consumers seeking to purchase used cars. The changing landscape of used vehicle markets makes it an ideal setting to understand the impact of online information on the value of certification – an issue that would be of considerable interest to academicians as well as practitioners.

We draw upon a unique and extensive dataset of consumers who obtained vehicle and transaction related information from online information sources in their used vehicle purchase process to examine the impact of their information acquisition on their choice of certified used cars, as well as the price paid. We compare the outcomes of sales where consumers purchased certified used cars with sales of used-cars where there was no certification, after controlling for a number of individual, and vehicle characteristics. Four different types of online information are found to be salient for buyers of used cars. Interestingly, each of them impacts the *value of certification* in different ways, as evidenced by their effects on consumers' choice (demand) and price. While prior empirical research has largely ignored the extent of information asymmetry in the markets for used goods, we show that the increased availability of different types of online information has a significant impact on consumers' choice of certified versus non-certified used cars as well as the price paid. A key implication then is that, as long as the cost of providing different categories of information is low, online infomediaries have important consequences for buyer and seller outcomes in traditional markets for used goods. We discuss the relevance of our findings for buyers and sellers in secondary markets, as well as for online infomediaries.

The rest of the paper is organized as follows. The following section describes the context of our study as well as the increasing importance of online information in the market for used cars. We then provide a brief overview of related research followed by a description of the data and measures used in the study. We then present the details of the empirical analyses, and the results of our study. We then discuss the findings and their managerial implications. The final section contains concluding remarks.

## Research Context

The used car market has been growing at a phenomenal pace. While 13.6 million new vehicles were purchased in North America in 2004, the corresponding numbers for used vehicles was 42.5 million—representing a 12% increase over 2003<sup>1</sup>. The used-car market is a classic example of a “lemons” market (Akerlof 1970) where sellers have private information about the quality of their offerings, while buyers are unable to assess quality with certainty. In order to overcome a potential market breakdown, sellers resort to different mechanisms to signal the quality of their products – certification, being the most common among them. Certification of used cars emerged as a byproduct of leasing in the late 1980s and 1990s, as luxury car manufacturers and dealers sought to resell the vehicles whose lease period had ended. *Certification* usually implies that the certified vehicle has been put through a comprehensive inspection process—anywhere from a 100-point to 300-point inspection and reconditioning<sup>2</sup>. Certified pre-owned vehicles have increasingly become an important category of vehicle purchases over the last decade. A J. D. Power and Associates study (2005) found that certified used cars accounted for 41% of all used-car dealership sales in 2005, up from 38% in 2004.

In the market for used cars, certification can potentially alleviate the problems of adverse selection by acting as a reliable signal of product quality. However, an interesting and crucial aspect of certification in this market is that, unlike situations where certification is generally provided by third-parties, certification of used cars is typically backed by the manufacturer, but ultimately provided by the sellers themselves (also known as first-party certification (Gereffi, Garcia-Johnson, and Sasser 2001)). Certified used cars typically sell for a premium over uncertified ones, and despite their growing popularity there have been several criticisms of these certification programs (Cutler 2005). Whether certified used vehicles offer value to consumers remains an

---

<sup>1</sup> Source: Manheim, Used Car Market Report, (Atlanta, GA), <http://www.manheimnews.com/UCMR/reports/ManheimUCMR2005-4uN73r/dt.3/3/05>.

<sup>2</sup> According to the National Automobile Dealers' Association (NADA), the term “certified” involves six categories of inspections: General evaluation; Under-hood evaluation; Exterior assessment; Interior evaluation; Required service and maintenance assessment; and Exterior detailing analysis.

empirical question that needs to be validated – a question that takes on added significance with the increasing availability of online information.

As noted earlier, the growth of the Internet and online information intermediaries has had a big impact on auto-retailing and how consumers shop for vehicles. Car buyers now have access to several resources including car-buying guides, car make-model and pricing information, pictures, reliability and safety information. According to a recent AutoShopper.com study by J. D. Power and Associates, 80% of used car buyers have access to the Internet, and 59% of the buyers used the Internet during their buying process in 2006. The growing popularity of online information is also demonstrated by the fact that in 2004, for the first time, more than twice as many late-model used vehicle buyers found the vehicle they purchased through the Internet than both newspaper and magazine classified ads combined (J. D. Power and Associates, 2006). Obtaining information can help consumers bridge their informational gaps by learning about the vehicles price across dealers, the used car details, its attributes (for e.g., known defects and issues for its vehicle class), its history, the available alternatives, as well as the details of certification, and warranties offered. Given that potential buyers can acquire relevant information from online sources to reduce the information asymmetry they face, it is important to assess whether such *online information* serves as a complement or substitute to the traditional certification mechanism common in secondary markets. However, the impact of these variables on consumer choices and market outcomes has not been examined before.

### **An Overview of Related Research**

*Certification:* A well-established stream of analytical research examines certification (Lizzeri 1999), and its role as an effective quality signaling mechanism. However, there exist very few empirical studies that have examined the outcomes of such mechanisms for firms and consumers. In markets with information asymmetry such as bond ratings, insurance, comics, sports cards trading, and used goods, certification has been shown to be a particularly valuable signal. In addition to helping reduce buyer's information asymmetry, certification

generates new information for all market players (Jin, Kato, and List 2006). By helping reduce search costs for customers seeking credible suppliers, certification also helps increase demand from buyers who otherwise may have not entered the market (Tirole 1988). Finally, certification can also help provide sellers with information on buyer's unobservable characteristics such as risk aversion (Ippolito and Mathios 1990; Jin and Leslie 2003). Prior empirical work finds mixed evidence of the effects of certification on market outcomes. Some studies document a positive relationship between certification and price and/or quality in thoroughbred racehorses (Wimmer and Chezum 2003), ISO9000 certified manufacturing programs (Terlaak and King 2006), collectible stamps (Dewan and Hsu 2004), and sports cards (Jin and Kato (forthcoming)). In contrast, a study (Kleiner and Kudrle 2000) examining the role of occupational licensing in dental health services, finds that while tougher licensing does raise prices, they do not improve outcomes for consumers. Most of this existing work focuses on certification by independent third-parties rather than first-party certification, which is the case in the market for used vehicles. In one of the few studies examining mandatory certification of used cars, Pratt and Hoffer (1985) find that in states that mandated "quality certification", the quality of vehicles traded were not significantly different than in those states without such laws. Ginter, Young and Dickson (1987) find that the depreciation, and also markup, is independent of reliability (quality) performance of vehicles in the used car market.

A few studies have also examined the role of alternate information-asymmetry reduction mechanisms in secondary markets. For instance, prior research has examined the impact of reputation on price-related outcomes, and broadly found a positive relationship between the reputation of the seller and the resulting price of the transaction (Resnick et al. 2006; Dewan and Hsu 2004). In a recent study, Dewally and Ederington (2006) compare different quality signals in the market for comic-books, and find that certification sends the strongest signal, with certified books averaging a premium of 50% over uncertified ones. The findings of these studies more generally suggest that when alternative information sources or signals are available in markets with information asymmetries, the diagnosticity of a quality-signaling mechanism such as certification depends on the valence of these other signals in the environment (Purohit and Srivastava 2001). While our study

controls for alternate information-asymmetry reduction mechanisms such as additional warranties, the primary focus of our study is on understanding the impact of online information availability on the value of certification in secondary markets. In contrast to the earlier studies that focus on the role of certification or price-related outcomes within secondary markets, our study is motivated by the belief that the dramatic increase in the volume and variety of online information calls for a re-examination of the value of certification to consumers.

*Information Search:* A number of studies in marketing and economics have examined consumers' information seeking behaviors (Beatty and Smith 1987; Kiel and Layton 1981, Moorman et al. 2004; Schmidt and Spreng 1996). Broadly, these studies have identified a number of factors including, buyer characteristics, situational, and environmental variables, as influencing buyers' information seeking behaviors. A few studies have also examined information search within the context of purchase of new automobiles (Furse, Punj and Stewart 1984; Klein and Ford 2003; Moorthy, Ratchford and Talukdar 1997; Punj and Staelin 1983; Ratchford, Lee and Talukdar 2003; Srinivasan and Ratchford 1993). While an extensive review of this research-stream is beyond the scope of our paper, what is most relevant to our study is research examining consumers' information search in secondary markets. Given the heightened nature of risk faced by buyers in these markets, information search has been found to be a dominant risk reduction strategy (Gabbott 1991). In examining the different types of uncertainties, Urbany, Dickson, and Wilkie (1989) find that information search is impacted by two types of uncertainties – *knowledge uncertainty* specific to a particular alternative, and *choice uncertainty* related to choice among competing alternatives, while Choudhuri (2000) finds that *performance risk*, and *financial risk*, affect information search as well. Researchers also find that buyers search and obtain more information from varied sources to reduce uncertainty related to the product, and the transaction prices (Cox 1967; Dowling and Staelin 1994; Srinivasan and Ratchford 1993).

More recently, researchers have sought to understand consumers' online information seeking behaviors (e.g., Johnson et al. 2004, Klein and Ford 2003), and the impact of online information on market outcomes (e.g., Hodkinson and Keil 2003, Zettelmeyer et al. 2005). Studies examining the impact of online information on

market outcomes find important differences in the outcomes resulting from the acquisition of price vis-à-vis product information (for instance, Diehl, Kornish and Lynch 2003; Lynch and Ariely 2000). Studies examining the impact of online price information find that providing price information leads to more competitive pricing by firms and lower average prices for consumers (Smith and Brynjolfsson 2001, Baye and Morgan 2001, Baye, Morgan and Scholten 2003, Chen, Iyer and Padmanabhan 2002). In contrast, research examining the role of product-related information finds that customers are less price-sensitive when provided with product quality information (Alba et al. 1997, Lynch and Ariely 2000). Together, these findings suggest that an examination of the search process must take into account the type of information obtained during the search process.

While most of these studies examine the impacts of online information on consumer outcomes in primary (new-product) markets, our study extends this stream of research by analyzing the impact of online information on outcomes in secondary markets. Moreover, much of the existing literature that examines the impacts of a focal information source (online information) on the choice and use of other media treats the Internet as an aggregate or monolithic source of information. However, it is clear that consumers can obtain several dimensions of online information related to purchasing a (used) product, each of which can have different impacts on outcomes in the market. Further, they fall short of assessing how online information interacts with alternate mechanisms such as certification to affect market outcomes. Our study seeks to fill these gaps.

## **Theoretical Background and Research Hypotheses**

We propose a simple research framework to investigate the impact of online information in secondary markets, particularly in the market for used cars, and posit hypotheses relating information acquisition to consumers' choice of certified vehicle purchase, and price paid.

### **Online Information Dimensions Salient to Used Vehicle Purchase**

As noted earlier, most of the research on the impact of online information has focused on the impact of *price-related* information on market outcomes. For instance, in the automobile industry, Zettelmeyer et al. (2005) find

that new vehicle buyers who use the Internet pay 2.2% less for their car than those who do not use it, to obtain a savings of \$500 on the average car. In the insurance industry, Brown and Goolsbee (2002) show that the growth of the Internet has reduced the price of term life insurance by 8%–15%. More recently, Zettelmeyer, Scott Morton and Silva-Risso (2006) demonstrate that the savings in car markets arise primarily through two means—by informing consumers of dealer invoice prices, and by providing buyers with online buying services that obtain attractive rates by aggregating demand across dispersed buyers.

In addition to *price information*, a few studies have also examined the impact of *product-related* information obtained by consumers from online sources. Ratchford et al. (2003) for instance, examine the role of the Internet as a source of product-related information, and how it impacts consumers' use of other conventional sources of information (such as TV and print) and their total search effort in the purchase of new automobiles. In a recent study of online auto-retailing, Viswanathan et al. (2007) find that consumers using online auto-retailing intermediaries obtain price and product related information, each of which has a differential impact on outcomes such as buyer's willingness to pay, market prices, and customer satisfaction. Other studies have also found that the availability of product information was found to increase buyers' willingness-to-pay by providing buyers with a product of better fit (Lynch and Ariely 2000); while access to price-information tended to reduce consumers' willingness-to-pay (Diehl et al. 2003).

Building on this past research, we construct a typology of information needs specifically relevant to the purchase of used goods. We map the product and price-related information dimensions identified in earlier research to product-features and transaction price related information. Product-features and transaction price information assume greater importance in the market for used goods as each used-good is unique and differentiated from others. In addition, compared to the new product purchases, consumers seeking to purchase used goods are faced with uncertainty about the *performance of the product* and the *value of their purchase*. Consequently, information that can help reduce uncertainty/risk related to the product quality (performance) and information that can help reduce uncertainty about the value (price) of the used good would be very valuable to

consumers. Since the purchase a used good is typically framed against the alternative of buying a new good, quality and price of the new good serve as critical reference points. Thus, the buyer needs to obtain (product and price) information not only about used goods, but also product and price-related information about new alternatives that may serve as close substitutes. These two additional information-needs lead us to propose a taxonomy consisting of four broad information needs specific to the used vehicle context.

Information buyers obtain from online search is therefore, categorized into product-related information including vehicle *features and vehicle reliability*, and price-related information on *transaction*, and *comparative alternatives*. Examining how each of these categories of online information substitutes or complements “information” provided by traditional certification forms a key focus of our study. In particular we are interested in examining if the acquisition of online information by consumers increases the *value of certification*. From an individual customer’s perspective, online information could lead to an increase/decrease in the likelihood of purchasing a certified used car, and/or increase/decrease the willingness-to-pay for a certified used car, compared to its uncertified counterpart. In other words, an increase in the *value of certification* could manifest as either an increase in the *demand for certified used cars* or an increase in the *price of certified used cars* or both. In the following section we develop hypotheses relating the impact of the four different types of online information to the choice of certification as well as price-related outcomes.

## **Hypotheses**

Certification acts as a signal for quality and therefore a buyer may be expected to be willing to pay a higher price for certified vehicles. On the other hand, certification in the case of used cars is provided by dealers themselves and hence might also be perceived as lacking in real value. The value of certification for used cars is thus an empirical question and one that we test in our first hypothesis, which posits that:

*H1. The price paid will be higher for purchase of a certified vehicle.*

In the market for used vehicles, certification serves mainly as a signal of quality, but can also provide information to the consumers about the condition of the used vehicle, including all repairs and maintenance

services that were performed upon the vehicle to guarantee a minimum standard of quality. As a signal of quality, certification is most closely associated with aiding in the reduction of uncertainty related to the quality and performance of the used vehicle. Reliability-related information helps to decrease performance uncertainty associated with the product (Kaplan, Szybillo, and Jacoby 1974), and is therefore expected to lower the value for certification. This lower value for certification could manifest itself as a lower demand for certified used cars and/or a lower price (or willingness to pay WTP) for a certified used car. Thus we posit that:

*H2a. Greater product Reliability information obtained by a buyer will reduce the likelihood of purchasing a certified vehicle.*

*H2b. Greater product Reliability information obtained by a buyer will decrease the price paid by the buyer for a certified vehicle.*

As above, product features-related information is likely to have a similar impact to reliability information on consumers' choice of certification and price paid. As described in Markopoulos et al. (2004), many markets face an information deficit, where in spite of the greater technological advancements and ease of providing information online, consumers continue to be imperfectly informed about product attributes due to the lack of sufficient information. Further, such information is shown to have key impacts on consumer outcomes by helping consumers gather better information about products before purchase, and also allowing them to efficiently search large and complicated product spaces according to their own personal needs and preferences, thereby mitigating knowledge uncertainty (Urbany et al. 1986). Thus, we expect that the retrieval of greater features-related information will reduce the buyer's sensitivity to quality issues and result in a lower value for certified used vehicles. As earlier, this lower value for certification could manifest itself as a lower demand for certified used cars and/or a lower price (WTP) for certified used cars. Thus we posit that:

*H3a. Greater product Features information obtained by a buyer will reduce the likelihood of purchasing a certified used vehicle.*

*H3b. Greater product Features information obtained by a buyer will decrease the price paid by the buyer for certified used vehicles.*

Armed with greater information about financing offers and discounts available for the vehicles, consumers are in a position to better understand the total financial costs of owning the car, reducing their financial/ economic uncertainty (Kaplan et al. 1974). Obtaining transactional price-related information may increase the consumers' perception of savings from making a purchase, and free up resources that can be invested in other affordable and useful add-ons, such as certification. In addition, consumers who obtain information about the costs of certification and additional warranties available for the vehicles obtain greater information about the value and benefits of certification. Hence, consumers who obtain transaction-related information are likely to perceive a greater value in certification. This higher value for certification could manifest itself as a higher demand for certified used cars and/or a higher price (WTP) for certified used cars. Thus we posit that:

*H4a. Greater Transaction price information obtained by a buyer will increase the likelihood of purchasing a certified vehicle.*

*H4b. Greater Transaction price information obtained by a buyer will increase the price paid by the buyer for a certified vehicle.*

Certified used vehicles are seen as a close substitute to new vehicles. In particular, certification is perceived to reduce the uncertainty associated with used vehicles making them closer to new vehicles. More importantly, certified used vehicles are seen as better alternatives to new vehicles as they are priced lower than new vehicles. Hence, consumers who obtain comparative information about new vehicle prices are likely to perceive lower choice uncertainty, as well as lower economic uncertainty about the value of their purchase, and consequently, greater value in certification. This higher value for certification could manifest itself as a higher demand for certified used cars and/or a higher price (WTP) for certified used cars. Thus we posit that:

*H5a. Greater Comparative price information obtained by a buyer will increase the likelihood of purchasing a certified vehicle.*

*H5b. Greater Comparative price information obtained by a buyer will increase the price paid by the buyer for a certified vehicle.*

## **Data and Measures**

### **Data**

Our study is based on secondary data obtained from a survey of 1999 to 2004 model-year used vehicle buyers, conducted by one of the largest market-research organizations in the US. The quota sampling strategy was designed to ensure that sufficient sample size was obtained for car make analysis, including a return of a minimum of 125 in-line<sup>3</sup>, certified returns for those car nameplates with certification programs, and a minimum of 120 responses from buyers who used the Internet for their automotive shopping for nameplates without certification programs. Two versions of an eight-page questionnaire (differentiated by rotating question items) and a cover letter along with a \$1 incentive were sent out in late January 2004 and a reminder postcard was mailed about a week later. Out of the total mail-out to a randomized sample of 78,534 buyers, 12,142 completed surveys were returned resulting in a response rate of 15.5%. The dataset consists of both consumers who used the Internet as part of their purchase process and traditional consumers (who did not use the Internet to shop for their used vehicle). Sampling weights based on a sales-weighting process are used to ensure that the distribution of makes in the sample was representative of the total personal use registrations of vehicles completed in the sampling period.

For the purpose of this study, we follow Zettermeyer et al. (2006) in defining a “car” as the “interaction of make, model, body type, transmission, displacement, doors, cylinders, and trim level” (p.170). To adequately control for vehicle fixed effects, we restrict our analysis to the top 135 Vehicle Identification Numbers (VINs), each of which had at least 25 responses or vehicle purchases. Thus our resulting sample consisted of a total of 5,349 vehicle purchases. We further restrict our sample to those make-model combinations which have both certified and non-certified sales in our data<sup>4</sup>. This restricted sample of 3,823 vehicles helps us achieve quasi-experimental control, and the various additional controls including vehicle and buyer characteristics among others, allows us to isolate the impact of online information.

---

<sup>3</sup> In-line refers to the fact that the dealer buyers bought from sold used and new vehicles of the same make.

<sup>4</sup> This reduction in the observations led us to drop 68 luxury and 419 non-luxury purchases of 10 “cars”, and observations with missing values. The average price of each segment rose slightly from \$23,205.36 to \$23,767.23 and from \$12,748.36 to \$12,788.98, respectively. The proportion of Internet users and online information found is not significantly different across the new and old sample. The restricted sample has a marginally lower proportion of certified purchases (33.09% vs. 33.78%).

57% of the consumers in our sample are male with median age of 51 years, while median age for females was 46 years. Racial minorities (African Americans and Hispanics) constitute 7.8% and buyers with low education (high school and below) constitute 24.0% of the sample. Buyers in our study have several options from where to purchase the used vehicle. A majority of the purchases in our sample were made at new car dealerships (63.12%), followed by used car dealers (20.73%), the rest being private sellers, and rental car companies. Our sample of respondents provided detailed information on their beliefs and use of used vehicle certification programs. 33.0 % purchased a certified used vehicle, and 3.8% had previously owned a certified used car. We obtain detailed information on buyers' online and offline information search processes. More than half of the buyers (56.3% of the sample) report using the Internet to help shop for their used vehicle, and spent an average of 6 hours doing so. 77% of the internet users conducted online research, on average, 6 weeks prior to visiting physical dealer locations. Those who visited online sites post dealer visits do so after 2 weeks, on average. Our sample of Internet buyers visited on average 3.38 third-party websites, 2.45 manufacturer websites (MFG), 3.24 dealership websites, 0.57 newspaper sites, and 0.09 chatroom/bulletin board sites.

## **Measures**

Table 1 provides a summary of the variables used in our empirical analyses. Our primary outcome variables are the buyer's choice of certification and the price paid for the used vehicle. *Certification* is measured as a binary variable, while *Price* is a continuous variable measured in dollars. The rest of the independent variables are measured as follows. *Information* buyers obtain from online search are categorized into product-related information about *vehicle reliability*, and *features*, and price-related information about *transaction* and *comparative options*. *Vehicle Reliability information* includes the historical performance of the make-model, vehicle reviews, safety and crash test ratings, and its road-handling abilities. *Vehicle features* includes information on specifications, access to vehicle photographs, and tools for comparing available features across different vehicle make-models. *Transaction information* includes information on special offers, discounts and financing options, along with warranty options and certification programs available on the vehicle that are of

interest to a potential used vehicle buyer. *Comparative information* includes prices of alternate options available to the potential buyer such as new vehicles and trade-in values. Buyers report about the various types of information found by them during their online search process. These are shown in Table 2. In order to extract the underlying information constructs, while reducing the number of parameters in our model, we employ Principal Component factor analysis with varimax rotation to obtain bundles of related information found by our buyers. Our four-factor model summarily captures the multi-faceted information structure (explains 74.61% variance in the items), and as shown in Table 2, exhibits reliability in an acceptable range from 0.68 to 0.72 for three of the factors, with Comparatives information displaying a marginal reliability of 0.59. Majority of the loadings are above 0.60 with the exception of tools for calculating monthly payments (0.53 on Transaction) and trade-in values (0.56 on Comparatives).

The dataset includes *vehicle characteristics* such as *Mileage*, *Model year*, and whether the vehicle was a *Luxury* make. We also include 125 dummies to represent the used-car make-model-trim feature combination as described earlier. The dataset also contains information on *buyer demographics* including *Gender*, *Age*, *Income*, *Low Education*, *Minority Race* and *Marital Status*. We control for buyer's purchase intentions using *Want Used*, a binary that indicates whether the buyer specifically only wanted to purchase a used vehicle, and buyer's experience with used cars- *Experience Used* - which is equal to 1 if the previously owned car was used.

Additional controls include buyer's access and use of offline information. We consider two such sources- use of Offline classifieds/ads in newspapers, TV and magazines (*Offline Ads*), and use of recommendations from friends, relatives and/or own personal experience to search and locate the seller/purchased vehicle (*Offline Personal*). We also control for the *characteristics of the market* and *the type of seller*. Market characteristics include dummies for *Metro*, *Small*, or *Rural* markets - the type of market where the car was purchased, and the seller type indicates whether the car was purchased from a *New car dealer*, a *Used car dealer*, or *Other* seller. Two variables allow us to control for the quality of the car in determining the effect of certification and information on the price paid. The first is a measure of buyer's overall *Satisfaction* with the

vehicle (scaled from 1 to 10), and the second, the *NumProblems* (number of problems) encountered in the vehicle within three months of purchase. These variables act as a proxy for the unobserved car quality. Finally, *Addwarranty*, equal to 1 if a buyer purchased any seller-provided additional warranty protection and 0 otherwise<sup>5</sup>. All interaction terms are formed after centering to reduce multicollinearity.

## Empirical Analyses

### Model specification

We specify the following econometric model to assess consumers' acquisition of the four types of online information, their choice of certification, and the price they pay for their used vehicle. The *price* equation is specified in [1], and the certification choice equation is given in [2]. The information choice equation is specified in [3] for the four different types of information (i =Features, Reliability, Transaction, Comparative).

$$\begin{aligned} \text{[1] PRICE} &= \alpha_1 + \beta_1 \text{VEHICLE} + \gamma_1 \text{BUYER} + \eta_1 \text{CERTIFICATION} + \sum \delta_{1i} \text{INFORMATION}_i + \theta_1 \text{INTERACTIONS} \\ &+ \zeta_1 \text{CONTROLS\_P} + \varphi_1 Z1 + \varepsilon_1 \end{aligned}$$

$$\text{[2] CERTIFICATION} = \alpha_2 + \beta_2 \text{VEHICLE} + \gamma_2 \text{BUYER} + \eta_2 \text{PRICE} + \sum \delta_{2i} \text{INFORMATION}_i + \zeta_2 \text{CONTROLS\_C} + \varphi_2 Z2 + \varepsilon_2$$

$$\text{[3] INFORMATION}_i = \alpha_{3i} + \beta_{3i} \text{VEHICLE} + \gamma_{3i} \text{BUYER} + \zeta_{3i} \text{CONTROLS\_I} + \varphi_{3i} Z3 + \varepsilon_{3i}$$

In equations [1] - [3], we include a vector of common *Vehicle and Buyer* variables as described above.

Certification is a binary variable indicating whether the buyer purchased a certified used vehicle. The

*Interactions* vector is a set of cross products between *Certification* and the four *Information* factors, and the two interactions of *Certification* with *Mileage* and *Luxury* make, to capture the differential impacts of certification across high mileage and luxury cars. Finally, we include several controls in the three equations. *Controls\_P* and *Controls\_C* includes *Offline Ads*, *Offline Personal*, *vehicle quality (Satisfaction, and NumProblems)*, *Market size (Small, Metro)*, *type of seller vehicle was purchased from (New car dealer, Used car dealer)*, and *Addwarranty*.

*Controls\_I* contains controls for offline information search- *Offline Ads and Offline Personal*. Since equations

---

<sup>5</sup> Additional warranty refers to any extended coverage (e.g., bumper to bumper or powertrain) purchased from the dealer/seller that is not included with certification.

[1]-[3] constitute a system of simultaneous equations, we also include a vector of instruments Z1-Z3 in each equation respectively to enable estimation of endogenous variables. These are discussed below.

### **Estimation Procedures**

The above system of equations highlights several concerns that arise in the process of estimating the impact of online information on the outcomes of interest.

*Treatment bias:* The first concern arises from the possibility of the existence of treatment bias (Heckman 1979), where non-random selection into treatment conditions (here, *the choice of certification*) in the sample leads to biased estimates if ignored. In our sample, treatment bias may arise from either demand or supply sources. It must be noted that selection effects caused by observables are accounted for by the 3SLS model. Therefore, here, we are concerned with selection caused by unobservables. Non-random self-selection of buyers that purchase certified vehicles occurs when unobserved variables that lead buyers to purchase certification also lead them to systematically obtain higher or lower prices. For instance, risk averse buyers would be more likely to buy certified used cars but also pay higher prices on an average (if they believe higher price to be a proxy for higher quality), compared to buyers with low risk-aversion. If this were the case, then the coefficient of *Certification* in equation [1] would be under-/over-estimated as it captures not only the effect of certification, but also of the correlated unobservables on price. We simultaneously estimate the price-certification equations using maximum likelihood estimation, along with an exclusion restriction (*Shop Certified*) that helps to robustly identify the model, and reduces our reliance on the functional form of the equations for identification. The results of a Likelihood ratio test suggest that treatment bias is not an issue for *Certification*.<sup>6</sup>

Second, there may be non-randomness in the vehicles that are chosen to be certified by the seller. Sellers may selectively choose to certificate certain types of vehicles (e.g., newer-model, low mileage, and luxury makes) that are more profitable to sell as certified, thereby upwardly biasing the coefficient of certification on

---

<sup>6</sup> These results are not shown here for sake of brevity, but are available from the authors.

price that is derived without accounting for selection of vehicles. Again, if selection is based on observables, 3SLS can be used to account for it. The coefficient between *Price* and *Certification* in our sample is 0.17( $p < 0.01$ ), suggesting either that higher value cars are certified, or alternatively, that certified cars sell for more. We deal with this issue by adding *Price* as an explanatory variable in [2].

*Error covariance:* A second concern relates to the possibility of contemporaneous error covariance across choice in [2] and [3] and price outcome in [1] for a given buyer, indicating that common unobservables influence all equations in the system, including information acquisition, choice of certification and price.<sup>7</sup> Another related issue arises from the presence of a common subset of right hand side regressors- *Vehicle* and *Buyer* characteristics. These problems could potentially confound the parameters of interest.

*Endogeneity:* Additionally, OLS assumptions may be violated due to the presence of first, non-recursive endogeneity between the choice of certification model and the outcome model, since certification and price are reciprocally causal (such that  $\varepsilon_2$  and  $\varepsilon_1$  are correlated), and second, endogeneity between online information and price, and certification (whereby  $\varepsilon_3$  is correlated with  $\varepsilon_2$  and  $\varepsilon_1$ ).<sup>8</sup> It is therefore necessary to account for the resulting confounding using appropriate instrumental variable estimation methods described in Maddala (1983) and Greene (2003).

We therefore utilize a full information estimation technique (3SLS) that takes into account both the effects of cross-equation *error covariance* and *endogeneity* by combining 2SLS and SUR.<sup>9</sup> In our 3SLS model, we

---

<sup>7</sup> Incidentally, this also addresses the problem of measurement error, if present, since each individual respondent provides multiple observations on use of online information, choice of certification, and price paid in our sample.

<sup>8</sup> This arises because whether the buyer finds online information is potentially an endogenous variable, correlated with unobservables in the price and certification equations. For example, certain more popular car make-models may also be associated with a larger number of websites that provide information about them, and a greater proportion of certified cars (made possible by the larger volume of vehicle trade), thereby increasing consumer's likelihood of finding related online information and purchasing certification.

<sup>9</sup> For comparison, we also estimate several other models: information, certification and price equations separately using probits and OLS, and simultaneously using SUR and 2SLS models. However, we only provide the results from the more complete 3SLS model.

treat certification and online information as endogenous in the price equation [1], and additionally, online information and price as endogenous in the certification equation [2]<sup>10</sup>.

*Identification:* We specify the following instrumental variables (Z2 and Z3) to help identify our system of equations. *Shop Certified* is a binary variable that describes that the buyer intentionally shopped only for a certified used vehicle. Following Wooldridge (2002), we also create cross products of the instruments of endogenous interaction terms for additional instruments. We also form interactions with *Shop Certified* and *Mileage* and *Luxury*. Further, we use as instruments buyers beliefs regarding certification programs—there is value in certification, certification programs offer special financing/discounts, and previous experience with certified programs. We include a dummy to indicate whether the car the buyer purchase just happened to be certified. We use an interaction between time spent on the Internet (hours/week) and *Shop Certified*. A set of four variables that measure the buyers' ratings of the importance of each type of information on a scale of 1 to 10, i.e. importance of FEAT, RELB, TRAN, and COMP, are used to instrument online information. Two variables are used to instrument for price- whether seller offered financing and whether buyer considered the vehicle purchased to be the best deal for the money.

The relevance of our instruments is confirmed by the significance of the F-test and the partial R<sup>2</sup> of the included set of instruments when regressed on the endogenous variables. Their validity is assessed as a test of the assumption of exogeneity between the set of instruments and the regression error of the outcome equation, conducted using tests of overidentification. The Sargan test statistic is not rejected in the model in Table 5, reinforcing the validity of our set of instruments and confirming that the set of instruments in each equation adequately identify our system of equations.

## Results

---

<sup>10</sup> This is further supported by the rejection of the Hausman test that there is no misspecification or difference in the coefficients obtained from the 2SLS and 3SLS models (chi-sq (153) = 697.26, p =0.00), thereby lending support for the latter.

Table 3 reports the summary statistics and Table 4 the correlations among constructs. Table 5 presents the main results of our analyses from 3SLS. M1a and M1b provide the coefficients for the estimation of log *Price* and *Certification* respectively, and M1c—M1f provides the coefficients of the estimation of buyers' retrieval of online information. In the following paragraphs we outline the results relating to the impact of each group of factors on the retrieval of online information (models M1c—M1f), choice of certification (denoted by model M1b), followed by their impact on price (denoted by model M1a).

### **Model Results**

The impacts of Vehicle Characteristics- *Mileage*, *Model year*, and 125 make-model dummies (not displayed) and vehicle *Price* are shown in Panel A; effects of Information, *Certification* and their interactions are in Panel B; Buyer demographics and psychographics are in Panel C; and Panel D includes offline information controls<sup>11</sup>.

*Vehicle Characteristics:* We find that most vehicle characteristics do not impact buyer's likelihood of obtaining online information. Interestingly though, we find that buyers purchased certification on higher mileage and early model (year) cars. As for the impact of vehicle characteristics on price, we find that the average buyer pays about \$100 less for each additional 1000 miles on the car, and an additional \$1098 for newer model cars.

*Certification and Online Information:* Our main interest lies in examining the effects of online information on the buyer's value for certification, as indicated by its impact on the likelihood of purchasing a certified used vehicle, and the price paid. Certification has a significant positive coefficient on price (supporting H1) indicating that buyers pay a premium (of about \$1158) for certified vehicles. While certified vehicles are more expensive, the average coefficient for certification covers the premium for luxury makes and lower mileage cars. This is evidenced in column M1a, where the coefficients of the interaction terms of mileage and luxury with certification are insignificant. In other words, after controlling for price in the certification equation, our results indicate that consumers do not pay a higher premium for certification on luxury and lower mileage cars.

---

<sup>11</sup> Other control variables described earlier, are included, but not displayed.

With respect to consumers' choice of certification and price paid for the used vehicle (as shown in Panel B, Table 5), we find that online information has a significant impact on a consumer's likelihood of purchasing a certified vehicle. Interestingly, the different categories of information (*product* and *price*) affect the choice outcome in different ways. Increased retrieval of product-related information such as *reliability* and *features information* from online sources *reduces* the likelihood of purchasing a certified vehicle. However, the retrieval of *transaction* price information and *comparative* price information *increases* the likelihood of a certified purchase. Thus, hypotheses H2a, H3a, H4a, and H5a are supported. With regards to price, we find that product and price related information have the opposite effects on price paid. Overall vehicle *reliability information* is found to have a positive impact on price, with consumers paying about \$627 more on an average, per unit of information obtained. The coefficient on features information, while positive, is however insignificant on price. Obtaining *transaction information* has a negative impact on price of nearly \$92 per unit of information, while comparative price information reduces price by \$208 per unit of information obtained.

Note that the coefficients of the interaction between certification and the four online information factors are all insignificant in M1a. This captures the marginal effect of online information on the price paid for certification. The split sample analysis in Table 6 further highlights the differential impact of online information on price paid across the certified and non-certified purchases. The results indicate that for certified cars, online information has no direct impact on price paid, indicating that hypotheses H2b-H5b are not supported. For non-certified cars, however, online information affects price in the same manner as in the pooled sample. Our results highlight that, while online information had no additional impact on price, beyond its impact on choice of certification for buyers who purchased certified vehicles; online information had significant (although, varied) effects on the price paid for non-certified vehicle purchases.

*Buyer Characteristics:* It is also informative to examine what kinds of buyers obtain online information. In panel C (Table 5), we present the results of the post-hoc analysis of buyer characteristics. We find that high income buyers obtain significantly more comparative price information, attesting to the fact that they are the

most likely group to consider buying new vehicles. Minority buyers are less likely to obtain all four types of online information. Buyers with low education are less likely to retrieve *features*, *reliability*, and *comparatives* information, and older buyers get less *features* information. Interestingly, gender and marital status play no role in differential acquisition of online information, choice of certification or price paid. With relation to buyer experience, those who previously owned used vehicles, obtained significantly less *reliability*, *transaction*, and *comparative* information, but are more likely to buy certified vehicles. Buyers who specifically wanted a used vehicle obtained more *features* information, but less *comparative* price information. Further, they were more likely to purchase certification.

Finally, we briefly consider the impacts of two additional information controls. Offline information search is an important option available to buyers, and we find that buyers who used offline classifieds and ads to locate their vehicles were also more likely to obtain online information, while buyers who relied on personal offline sources were less likely to obtain online information. Both types of offline sources were associated with increased propensity of purchasing certification. However, only information from the latter—personal sources—helps to significantly lower the price paid.

### **Limitations**

Prior to discussing the implications of our findings it is important to acknowledge the limitations of this work. First, we are limited by our reliance on secondary data collected by a third party. However, this detailed data set collected by one of the largest market research firms in the US represents one of the most extensive surveys of used vehicle buyers and the measures used possess good psychometric properties. Second, there is a possibility of common methods bias as the data was collected through a single survey instrument. This is mitigated to a large extent by having each response correspond to vehicle registrations and tied objectively to a verified purchase. Additionally, we can reject the presence of this bias on the basis of Harman's single factor test (see Podsakoff et al. 2003 for a complete discussion) on the perceptual indicators, where we obtained multiple factors with eigenvalues greater than 1, with the largest factor accounting for less than 25% of the

variance. Third, our method of estimation using 3SLS offers some advantages- as a full-information model, it is theoretically more efficient because it uses information in the data more comprehensively. However, it potentially suffers from biases arising from model misspecification. To assess robustness, we also additionally estimate limited information models such as 2SLS, which do not face the same problem since they ignore information about the joint distribution of the error terms across the equations, albeit suffering from a potential loss of efficiency, if in fact 3SLS was appropriate. The results of these additional tests are robust to those of the 3SLS estimation, with no variables changing sign or significance. The consistency of the results across these different models reinforces the validity of our findings. Finally, it should be noted that one of our dependent variable – certification - is binary. Therefore employing 3SLS may lead to possible biases due to the dependency of errors on the coefficients. While there is no easily available technique to estimate structural systems of equations with binary endogenous variables, Angrist and Krueger (2001) show that this is not necessarily a perilous problem since the consistency of the second-stage or outcome estimates does not depend on getting the functional form of the first-stage or choice equation correct, and therefore linear regression is sufficient. In addition, to assess robustness, we also ran our models using probit regressions for first-stage choice models as well, and obtain consistent results.

## **Discussion**

### **Implications**

Certified pre-owned programs help manufacturers keep used-car residual values high and create vehicles with higher resale values. Certified used cars are also believed to be more profitable to dealers. Consequently, manufacturers as well as dealers have a strong incentive to promote certified used cars. As for consumers, certification may increase aggregate consumer surplus by increasing the average quality of cars traded in the used vehicles market. Certification also expands the market by making luxury brand vehicles affordable to consumers that would have otherwise not been able to purchase them. However, since such certification is done by the manufacturers/dealers themselves, the value of such certification to consumers has been

questioned. The presence of alternate mechanisms such as warranties, add to this debate about the value of certification in the market for used cars. Our findings show that even after accounting for the impacts of online information, and additional warranties, certified cars commanded a premium, suggesting that consumers have a positive valuation for certification. The figures in Table 4 highlight significant differences between the population of certified vehicles and non-certified ones. For instance, certified vehicles were more likely to be low mileage with lower variance in usage ( $t=7.85$ ,  $p<0.01$ ), newer model year ( $t = 6.50$ ,  $p<0.01$ ), and luxury makes ( $t=8.07$ ,  $p<0.01$ ). This suggests that consumers might benefit from this selective culling of certified used cars, which might be particularly valuable to risk-averse consumers as it allows them to enter the market for used vehicles. Thus buyers who might otherwise not consider purchasing a used vehicle might be able to purchase a certified used vehicle (Vella 2006)<sup>12</sup>. We also find that consumers who were more satisfied with their certified vehicle were significantly more likely to recommend their make-model to others (82.25% vs. 72.35%), and purchase a new vehicle of same make in the future than others (49.14% vs. 39.83%). In addition, consumers who bought certified cars were more likely to return to their dealer for post-purchase services than those who didn't (79.35% vs. 60.12%). These highlight the additional benefits dealers/manufacturers get from certification.

While certification, as seen above, plays an important role in the used vehicle market, one of the most significant developments in auto-retailing has been the dramatic increase in the amount and variety of online information available to consumers. For instance, consumers now have easy access to detailed information on crash tests from insurance companies, results of dependability surveys done by Consumer Reports, safety and reliability information from sites such as [www.AutoSafety.org](http://www.AutoSafety.org) and [www.LemonaidCars.com](http://www.LemonaidCars.com), as well as information about the used-car's history from services like CarFax.com. In addition, online information intermediaries such as KBB.com, Edmunds.com, Cars.com, IntelliChoice.com, and AutoTrader.com, among others, provide ratings, reviews, rankings on several reliability dimensions, and a plethora of information of

---

<sup>12</sup> The risk aversion may arise due to several factors including from the consumer lacking sufficient knowledge about cars, repair and maintenance.

value to consumers of used cars. However, their impact on consumer behavior and choices has not been examined before. We find that, after controlling for detailed vehicle and buyer characteristics, buyer pre-purchase vehicle consideration sets, as well as offline information sources, buyers' *value for certification* is significantly impacted by information retrieved from online sources. However, this impact depends on the type of information obtained by consumers (see Fig. 1). While information about used vehicle features can help to reduce the knowledge uncertainty for consumers seeking used cars, information regarding vehicle reliability/safety including data on vehicle speed, handling, and road performance, can help reduce performance uncertainty. We find that these two categories of information – *features* and *reliability* information – reduce the likelihood of a buyer purchasing a certified used car. In addition, we find that of the two product-related information categories, reliability information significantly increases the price paid by consumers purchasing non-certified used cars. Thus, from a seller's perspective reliability information is a stronger substitute to traditional certification, while complementing the sale of non-certified used cars<sup>13</sup>. On the other hand, transaction price related information can serve to reduce the economic/financial uncertainty, while comparative information relating to the price of new alternatives can help reduce choice uncertainty as well as economic uncertainty about the value of the purchase. In contrast to the impact of product-related information categories, we find that *transaction* and *comparative* price information - are *complementary* to the purchase of certification, while reducing the price paid by consumers purchasing non-certified used car purchases.

	<b>Features Information</b>	<b>Transaction Information</b>	<b>Reliability Information</b>	<b>Comparative Information</b>
<b>Certification Choice</b>	–	+	–	+
<b>Non-Certified Price</b>	+ (N.S.)	–	+	–

**Fig 1: Impact of Online Information on Choice of Certification and Price of Used Cars**

<sup>13</sup> It is pertinent to note that since our dataset consists of only two categories of used cars – certified and non-certified used cars - information that decreases the likelihood of a certified purchase increases the likelihood of a non-certified purchase.

These findings have significant implications for dealers of used cars. Given that used cars are twice as profitable for dealers as new vehicles (CIRP 2007), understanding the impact of the Internet on consumer outcomes becomes paramount. Price-related information not only increases a consumer's likelihood of purchasing a certified used car but also decreases the price paid for consumers purchasing non-certified used cars. Product-related information has the exact opposite impact on *consumers' choices* (and consequently the sellers' demand) and *price*. Understanding the differential impacts of these information categories on consumer choice and price outcomes is crucial for traditional dealers. Sellers of certified used cars would benefit from providing consumers with easier and inexpensive access to price and transaction related information, while sellers of non-certified used cars would benefit from providing consumers with greater access to information about product features and reliability.

Our findings also have implications for the strategic partnerships between traditional dealers and online information providers. Since online information providers vary in the type of information they provide, partnering with the right online information providers would be mutually beneficial to dealers as well as online infomediaries. For instance, dealers of certified used cars would benefit from affiliations with online infomediaries such as Capital One Auto Finance, and E-loans.com that provide transaction information. In addition, they would also benefit from partnering with online sites that provide information on *both* used and new vehicles. As indicated by our results, buyers who obtained comparative information on new car alternatives were more likely to purchase certification, highlighting the interrelationships between new and certified-used car sales. Alternatively, dealers selling non-certified used cars (for instance, dealers of non-luxury and high-mileage used cars), as well as consumers would benefit from dealers' affiliation with online infomediaries such as AutoSafety.org, LemonaidCars.com, and CarFax.com that provide reliability and safety information that serve as a substitute to certification but complement non-certified used car purchases. Identifying the right online information partners would not only help traditional dealers target the right customer segments but also

optimize their inventory of certified and non-certified used cars. Concomitantly, online infomediaries would also benefit by better highlighting the value of their information in reducing asymmetries in such markets.

Our findings relating to buyer characteristics and their likelihood of obtaining online information have some interesting implications. A recent survey conducted by Automotive Retailing Today, a coalition of automakers and dealers whose stated goal is to narrow the gap between media accounts of dealership conditions and consumers' experiences, finds that the majority of the minority buyers that were surveyed said that their dealership did not give them enough information to make an informed purchase, and that the dealerships often did not honor their commitments (Harris 2005). Our results indicate that minorities as well as less-educated consumers are also less likely to obtain the various categories of online information prior to their purchase. This has important implications for their welfare, as these are typically the consumers who tend to be discriminated against by traditional dealers as evidenced by past research findings (Ayres and Seigelman 1995; Scott-Morton, Zettelmeyer, and Silva-Risso 2003). Online information intermediaries can add greater value to these consumers who are more prone to discrimination in traditional channels. Currently, "the "Car Buyer's Bill of Rights" forces dealers to reveal vehicle history along with a copy of the inspection report when selling certified used vehicles and provide a two-day sales contract cancellation policy" (CIRP, 2007). However, our findings suggest the need for stronger public policy measures to ensure greater transparency in transactions with "disadvantaged consumers".

Our findings relating to the impact of online information on price of used cars also extend the earlier findings. For instance, Zettelmeyer et al. (2005, 2006), find that online buyers paid on an average about 2% less than offline buyers. In our study, by teasing out different types of online information sought by used vehicle buyers, we obtain more nuanced effects of online channel use. Another interesting finding relates to consumers' use of online and offline sources of information. We find that while impersonal/ commercial sources of offline information (for e.g., classifieds in TV/magazines/radio) complement online information search, both prior experiences as well as the use of personal information sources (for e.g., friends and relatives) serves as a

substitute to online information search in the context for used cars – a likely indication of the importance of trust in the decision to purchase used goods. This suggests that dealers of used goods might benefit from cost-effective alternate quality signals such as reputation mechanisms and ratings from earlier transactions to engender greater trust in consumers.

## **Conclusion**

Secondary markets are an important part of the economy and have been growing rapidly in many product categories. Clearly, secondary markets are an important category for vehicle manufacturers and play an important role in the demand as well as the profitability of new cars for manufacturers, as well as dealers. The rapid growth of the Internet and online information sources has dramatically changed the balance of power between consumers and car dealers. While earlier studies have examined the impact of the Internet on the market for new cars, there have been very few studies of secondary markets in general, and more specifically the impact of the Internet on the market for used cars. Our study is among the first to examine the impact of different types of online information on the market for used cars. In addition, we focus on the impact of such information on consumers' value for certification for used cars. Our study highlights interesting relationships between different types of online information and consumers' value for certification. Our findings relating to the impact of various information categories highlight the need to disentangle these effects empirically to better understand their differential impacts on the outcomes of interest to buyers and sellers. In general, our findings about the impact of the different types of online information on consumer demand for traditional quality signals and price outcomes provide useful guidelines for other secondary markets.

While the primary focus of this study has been on impact of online information on certification in the market for used cars, other mechanisms such as seller quality, warranties, product guarantees, branding, etc. also serve to reduce information asymmetries in many markets. It would be useful to examine the impact of the increased availability of online information on alternate quality signaling mechanisms. In addition, while this study examines the impact of information on consumer choices in a classic lemons market, it would be

interesting to study the impact of such information on the quality signaling mechanisms in online secondary markets such as eBay and Amazon, which have gained prominence.

## References

- Akerlof, George A. (1970), "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84(3), 488-500.
- Alba, Joseph, Joseph Alba, John Lynch, Barton Weitz, Chris Janiszewski, Richard Lutz, Alan Sawyer, and Stacy Wood (1997), "Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces," *Journal of Marketing*, 61(July), 38-53.
- Angrist, Joshua D. and Alan B. Krueger (2001), "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives*, 15(4), 69-85.
- Ayres, Ian and Peter Siegelman (1995), "Race and Gender Discrimination in Bargaining for a New Car," *American Economic Review*, 85(3), 304-322.
- Baye, Michael R. and John Morgan (2001), "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets," *American Economic Review*, 91(3), 454-474
- Baye, Michael R., John Morgan, and Patrick Scholten (2003), "The Value of Information in an Online Consumer Electronics Market", *Journal of Public Policy & Marketing*, 22(1), 17-25.
- Beatty, S. E. and Smith, S. M. (1987), "External search Effort: An Investigation Across Several Product categories," *Journal of Consumer Research*, 14(June), 83-95.
- Brown, Jeffrey R. and Austan Goolsbee (2002), "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry," *Journal of Political Economy*, 110(3), 481-507.
- Chaudhuri, Arjun (2000), "A Macro Analysis of the Relationship Between Product Involvement and Information Search: The Role of Risk, *Journal of Marketing Theory and Practice*, Winter 2000, 1-15.
- Chen, Yuxin, Ganesh Iyer, and V. Padmanabhan (2002), "Referral Informediaries," *Marketing Science*, 21(4), 412-434.
- CIRP (2007). "The Anatomy and Physiology of the Used Car Business". Research Report. Car Internet Research Program II.
- Cutler, Kim-Mai. (2005), "Certified Used Cars Come Under Fire: Wave of Lawsuits Claim Widely Offered Guarantees Fail to Disclose Past Problems," *The Wall Street Journal* (July 12).
- Cox, D. F. (1967), *Risk taking and information handling in consumer behavior*, Harvard University Press, Cambridge, MA.

- Dewally, Michael and Louis Ederington (2006), "Reputation, Certification, Warranties, and Information as Remedies for Seller-Buyer Information Asymmetries: Lessons from the Online Comic Book Market," *Journal of Business*, 79, 693-729.
- Dewan, Sanjeev and Vernon Hsu (2004), "Adverse Selection in Electronic Markets: Evidence from Online Stamp Auctions," *Journal of Industrial Economics*, 52(4), 463-590.
- Diehl, Kristin, Laura J. Kornish and John G. Lynch, Jr. (2003), "Smart Agents: When Lower Search Costs for Quality Information Increase Price Sensitivity," *Journal of Consumer Research*, 30(1), 56-71
- Dowling, Grahame R. and Staelin, Richard (1994), "A Model of Perceived Risk and Intended Risk-Handling Activity," *Journal of Consumer Research*, 21(1), 119-134.
- Folta, Timothy B. and Jay J. Janney (2004), "Strategic benefits to firms issuing private equity placements," *Strategic Management Journal*, 25(3), 223-242.
- Furse, D. H., G. H. Punj, and D. W. Stewart (1984), "A Typology of Individual Search Strategies Among Purchasers of New Automobiles," *Journal of Consumer Research*, 10(March), 417-431.
- Gabbott, Mark (1991), "The Role of Product Cues in Assessing Risk in Second-hand Markets," *European Journal of Marketing*, 25(9), 38-50.
- Gereffi, Gary, Ronie Garcia-Johnson and Erika Sasser (2001), "The NGO-Industrial Complex," *Foreign Policy*, 125, 56-66.
- Ginter, James L., Murray A. Young and Peter R. Dickson (1987), "A Market Efficiency Study of Used Car Reliability and Prices," *The Journal of Consumer Affairs*, 21(2), 258-276.
- Greene, William (2003), *Econometric Analysis*, 5th Edition, Prentice Hall.
- Harris, Donna (2005). "Poll: Buyers are unhappy with dealers," *Automotive News*, 79, (January 31).
- Heckman, James (1979), "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153-161.
- Hodkinson, Christopher and Geoffrey Keil (2003), "Understanding Web Information Search Behavior: An Exploratory Model," *Journal of End User Computing*, 15(4), 27-48.
- Ippolito, Pauline M. and Alan D. Mathios (1990), "Information, Advertising and Health Choices," *RAND Journal of Economics*, 21(3), 459-480.
- J. D. Power and Associates (2005), *Internet Overtakes Print Advertising as Primary Shopping Source for Used Vehicles*. Management Research Report (June 16). Accessed at <http://www.jdpower.com/corporate/news/releases/pressrelease.asp?ID=2005080>

- J. D. Power and Associates (2006), *Autoshopper.com Study: Online Used-Vehicle Classifieds Outpace Print Ads Two to One (May 25)*. Accessed at <http://www.jdpower.com/corporate/news/releases/pressrelease.asp?ID=2006076>
- Jin, Ginger Z. and Andrew Kato (Forthcoming), "Price, Quality and Reputation: Evidence from an Online Field Experiment," *RAND Journal of Economics*.
- Jin, Ginger Z. and Phillip Leslie (2003), "The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics*, 118(2), 409-451.
- Jin, Ginger Z., Andrew Kato and John A. List (2006), "That's News to Me! Information Revelation in Professional Certification Markets," Working Paper, University of Maryland, College Park.
- Johnson, Eric J., Wendy W. Moe, Peter S. Fader, Steven Bellman, Gerald L. Lohse (2004), "On the Depth and Dynamics of Online Search Behaviors," *Management Science*, 50(3), 299-308.
- Kaplan, Leon B., George J. Szybillo, and Jacob Jacoby (1974), "Components of Perceived Risk in Product Purchase: A Cross-Validation," *Journal of Applied Psychology*, 59(3), 287-291.
- Keil, Geoffrey C., and Roger A. Layton (1981), "Dimensions of Consumer Information Seeking Behavior," *Journal of Marketing Research*, 18(May), 233-239.
- Klein, Lisa R. and Gary T. Ford (2003), "Consumer Search for Information in the Digital Age: An Empirical Study of Prepurchase Search for Automobiles," *Journal of Interactive Marketing*, 17(3), 29-49.
- Kleiner, Morris M., and Robert T. Kudrle (2000), "Does Regulation Affect Economic Outcomes? The Case of Dentistry," *Journal of Law and Economics*, 43(2), 547-582.
- Lizzeri, Alessandro (1999), "Information Revelation and Certification Intermediaries," *Rand Journal of Economics*, 30(2) 214-231.
- Lynch, Jr. John. G., and Dan Ariely (2000), "Wine Online: Search Costs Affect Competition on Price, Quality and Distribution," *Marketing Science*, 19(1), 83-104.
- Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge.
- Moorman, Christine, Kristin Diehl, David Brinberg, and Blair Kidwell (2004), "Subjective Knowledge, Search Locations, and Consumer Choice," *Journal of Consumer Research*, 31(3), 673-680.
- Moorthy, Sridhar, Brian T. Ratchford, and Debabrata Talukdar (1997), "Consumer Information Search Revisited: Theory and Empirical Analysis," *Journal of Consumer Research*, 23(4), 263-277.

- Podsakoff, Phillip M., Scott B. Mackenzie, Jeong-Yeon Lee, and Nathan P. Podsakoff (2003), "Common Methods Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology*, 88(5), 879-903.
- Pratt, Michael D. and George E. Hoffer (1985), "State Used Motor Vehicle Disclosure Laws: Do They Make a Difference to the Consumer?" *Journal of Consumer Affairs*, 19(1), 177-185.
- Punj, Girish N., and Staelin, Richard. (1983), "A Model of Consumer Information Search Behavior for New Automobiles," *Journal of Consumer Research*, 9(March), 366-380
- Purohit, Devavrat and Joydeep Srivastava (2001), "Effect of Manufacturer Reputation, Retailer Reputation, and Product Warranty on Consumer Judgments of Product Quality: A Cue Diagnosticity Framework," *Journal of Consumer Psychology*, 10(3), 123-134.
- Ratchford, Brian T., Debabrata Talukdar and Myung-Soo Lee (2001), "A Model of Consumer Choice of the Internet as an Information Source," *International Journal of Electronic Commerce*, 40(2), 193-209.
- Ratchford, Brian T., Myung-Soo Lee, and Debabrata Talukdar (2003), "The Impact of the Internet on Information Search for Automobiles," *Journal of Marketing Research*, 40(2), 193-209.
- Ratchford, Brian T., Myung-Soo Lee, and Debabrata Talukdar (2007), "The Impact of the Internet on Information Search for Automobiles: A Re-Inquiry," *Journal of Consumer Research*, 34(1), 111-119.
- Resnick, Paul, Richard Zeckhauser, John Swanson, and Kate Lockwood (2006), "The Value of Reputation on eBay: A Controlled Experiment," *Experimental Economics*, 9(2), 79-101.
- Schmidt, J. B. and R. A. Spreng (1966), "A Proposed Model of External Consumer Information Search," *Journal of the Academy of Marketing Science*, 24(3), 246-256.
- Scott-Morton, Fiona, Florian Zettelmeyer and Jorge Silva-Risso (2003), "Consumer Information and Price Discrimination: Does the Internet Affect the Pricing of New Cars to Women and Minorities?" *Quantitative Marketing and Economics*, 1(1), 65-92.
- Smith, Michael D. and Erik Brynjolfsson (2001), "Consumer Decision-Making at an Internet Shopbot: Brand Still Matters," *Journal of Industrial Economics*, 49(4), 541-558.
- Srinivasan, Narasimhan and Brian T. Ratchford (1991), "An Empirical Test of a Model of External Search for Automobiles," *Journal of Consumer Research*, 18(September), 233-242.
- Terlaak, Ann and Andrew A. King (2006), "The effect of certification with the ISO 9000 Quality Management Standard: A signaling approach," *Journal of Economic Behavior and Organization*, 60(4), 579-602.
- Tirole, Jean (1988), *The Theory of Industrial Organization*, MIT Press, Cambridge, MA.

- Urbany, Joel. E., P. R. Dickson, and W. L. Wilkie (1989), "Buyer uncertainty and information search," *Journal of Consumer Research*, 16, 208–215.
- Vella, Matt (2006), *Best Used Cars*. Business Week Online (July 6).
- Viswanathan, Siva, Jason N. Kuruzovich, Sanjay Gosain, and Ritu Agarwal (2007), "Online Infomediaries and Price Discrimination: Evidence from the Auto-Retailing Sector," *Journal of Marketing*, 71(3), 98-107.
- Wimmer, Bradley S., and Brian Chezum (2003), "An Empirical Examination of Quality Certification in a Lemons Market," *Economic Inquiry*, 41(2), 279-291.
- Wooldridge, Jeffrey (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press.
- Zettelmeyer, Florian, Fiona Scott Morton, and Jorge Silva-Risso (2005), "Cowboys or Cowards: Why Are Internet Car Prices Lower?" *Discussion paper, Haas School of Business, University of California, Berkeley*.
- Zettelmeyer, Florian, Fiona Scott Morton, Jorge Silva-Risso (2006), "How the Internet Lowers Prices: Evidence from Matched Survey and Auto Transaction Data," *Journal of Marketing Research*, 43(2), 168-181.

## Tables and Figures

**Table 1**  
**Operationalization of Constructs**

<b>Construct</b>	<b>Operationalization</b>
Vehicle Price	Total Price in \$ (excl. tax, license, trade-in)
Model Year	Vehicle Model Year (1999 upto 2004) with base year 1998
Income	15 point scale (0 – Less than \$25K; 15 - More than \$250K)
Minority Race	1 – If African American or Hispanic; 0 – Otherwise
Low Education	1 – If less than a high school graduate, 0 – If more educated
Gender	1 – Male; 0 – Female
Age	Age in Years
Married	Whether married or not
Previous car was used	1 – Has previously owned a used car; 0- otherwise
Want Used	1 – Intended to purchase used vehicle; 0 – otherwise
Certified Vehicle	1 – Purchased certified vehicle; 0 – Otherwise
Mileage	Miles at Purchase /1000
Luxury Make	1 – If Luxury nameplate; 0 – Otherwise
Features Information FEAT	Average of 3 items corresponding to online information found on a scale of 0-1
Reliability Information RELB	Average of 4 items corresponding to online information found on a scale of 0-1
Transaction Information TRAN	Average of 4 items corresponding to online information found on a scale of 0-1
Comparatives Information COMP	Average of 3 items corresponding to online information found on a scale of 0-1
Visit Online Before Dealer	1 - conducted online search prior to visiting dealers; 0- otherwise
Use offline classifieds/ads	Average of 2 items corresponding to offline classifieds/ads used to locate and research vehicle 0-1
Use offline personal sources	Average of 2 items corresponding to prior experience and recommendations from friends/relatives used to locate and research vehicle 0-1
Additional Warranty	1 – Purchase additional warranty; 0 – Otherwise
Satisfaction with vehicle overall quality (Rating)	10 point scale for overall rating of vehicle
Post-Purchase Defects	Number of problems encountered with vehicle after purchase
Market location	Dummies for Rural, small, metro

**Table 2**  
**Factors for Online Information Search**

Please tell us whether you found this information while searching on the Internet (Yes/ No)

<b>Information Found While Searching Online</b>	<b>F1:Features</b>	<b>F2:Reliability</b>	<b>F3:Transaction</b>	<b>F4:Comparatives</b>
Used vehicle photographs	<b>0.7113</b>	0.2042	0.1424	0.0422
Locate used vehicles for sale	<b>0.6901</b>	0.0658	0.2290	-0.0005
Options and features available on used vehicles	<b>0.6044</b>	0.3672	0.1632	0.1405
Road tests/articles about vehicles	0.1165	<b>0.8192</b>	0.0961	0.1084
Performance data on vehicles (speed, handling, etc.)	0.1924	<b>0.7957</b>	0.1107	0.1033
Vehicle reliability information	0.0370	<b>0.7539</b>	0.1260	0.1887
Safety information	0.1291	<b>0.7274</b>	0.2315	0.0519
Special financing/discount offers	0.0909	0.1186	<b>0.7648</b>	0.2008
Service contract/extended warranty information	0.0499	0.1845	<b>0.7296</b>	0.1142
Information on certified used vehicles	0.1971	0.2197	<b>0.6811</b>	0.0232
Tool for calculating monthly payments	0.2851	0.1220	<b>0.5289</b>	0.1916
Dealer cost/invoice of new vehicles	-0.0255	0.1522	0.3027	<b>0.7229</b>
Prices of new vehicles	0.2890	0.1360	0.0786	<b>0.6823</b>
Trade-in values	0.0636	0.3045	0.1161	<b>0.5580</b>
<b>Factor Reliability (Cronbach's alpha)</b>	<b>0.68</b>	<b>0.83</b>	<b>0.72</b>	<b>0.59</b>
F1:Features	1.00	0.3922***	.3431***	0.3657***
F2:Reliability		1.00	0.3646***	0.3824***
F3:Transaction			1.00	0.3510***
F4:Comparatives				1.00

**Notes:** \*\*\* indicates correlation significance at p<0.01 level

**Table 3**  
**Split sample t-tests**

<b>Variables</b>	<b>Non-Certified Vehicles (N =2597)</b>	<b>Certified Vehicles (N = 1226)</b>	<b>t- Statistics (df = 3821)</b>
<b>Vehicle Characteristics</b>			
Price paid (\$1000)	14.57 (7.44)	17.37 (7.90)	t = -11.5917(p = 0.00)
Miles on Vehicle (000's)	36.65 (21.67)	31.08 (17.60)	t = 7.8508 (p = 0.00)
Luxury Import	0.21 (0.41)	0.33 (0.47)	t = -8.0673 (p = 0.00)
Model year	2.78 (1.34)	3.08 (1.32)	t = -6.5015 (p = 0.00)
<b>Consumer Demographics</b>			
Age	47.40 (14.85)	49.65 (14.87)	t= -4.3681 (p = 0.00)
Gender	0.59 (0.49)	0.55 (0.50)	t= 2.3572 (p = 0.02)
Low Education	0.24 (0.43)	0.23 (0.42)	t = 0.5665 (p = 0.57)
Income	6.05 (3.42)	6.33 (3.50)	t = -2.4043 (p = 0.02)
Minority Race	0.07 (0.25)	0.10 (0.30)	t = -3.3595 (p = 0.00)
Married	0.70 (0.46)	0.70 (0.46)	t = -0.0775 (p = 0.94)
<b>Consumer Experience/ Psychographics</b>			
Previous car was used	0.62 (0.49)	0.58 (0.49)	t = 2.2656 (p = 0.02)
Want Used vehicle	0.65 (0.48)	0.58 (0.49)	t = 4.1669 (p = 0.00)
<b>Online Information</b>			
Reliability Information	0.35 (0.40)	0.39 (0.41)	t = -2.5240 (p = 0.01)
Features Information	0.47 (0.45)	0.49 (0.44)	t = -1.6974 (p = 0.09)
Transaction Information	0.18 (0.28)	0.24 (0.32)	t = -5.8493 (p = 0.00)
Comparatives Information	0.36 (0.39)	0.40 (0.40)	t = -2.6801 (p = 0.01)
<b>Controls</b>			
Offline Classifieds/ads	0.55 (1.19)	0.50 (1.15)	t = 1.4752 (p = 0.00)
Offline Personal	0.34 (1.31)	0.30 (1.25)	t = 1.1291 (p = 0.00)
Bought from rural market	0.41 (0.49)	0.33 (0.47)	t = 4.5940 (p = 0.00)
Bought from small market	0.43 (0.49)	0.48 (0.50)	t = -3.0551 (p = 0.00)
Bought from new vehicle seller	0.57 (0.50)	0.77 (0.42)	t = -14.1365 (p = 0.00)
Bought from used vehicle seller	0.24 (0.43)	0.14 (0.35)	t = 7.6818 (p = 0.00)
Additional warranty	0.22 (0.41)	0.31 (0.46)	t = -6.7980 (p = 0.00)

Notes: \* p <.10, \*\* p < 0.05, \*\*\* p < 0.01; unpaired sample t-tests. Standard error of mean deviation shown in parentheses

**Table 4**  
**Correlations (N = 3823)**

	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 Price (1000)	15.57	7.77	1.00																
2 Miles	34.42	20.38	-0.34 ***	1.00															
3 Model year	2.90	1.35	0.22 ***	-0.60 ***	1.00														
4 Luxury	0.25	0.43	0.61 ***	-0.04 ***	-0.11 ***	1.00													
5 Income	6.10	3.46	0.33 ***	-0.10 ***	-0.04 ***	0.39 ***	1.00												
6 Minority	0.08	0.27	-0.00	0.04 ***	-0.02 *	0.01	-0.07 ***	1.00											
7 Education (Low)	0.24	0.43	-0.12 ***	0.03 **	0.06 ***	-0.17 ***	-0.26 ***	-0.00	1.00										
8 Gender (Male)	0.57	0.49	0.09 ***	0.02	-0.04 **	0.10 ***	0.18 ***	-0.06 ***	-0.06 ***	1.00									
9 Age	48.91	14.96	0.13 ***	-0.16 ***	0.20 ***	0.08 ***	-0.01	-0.09 ***	0.16 ***	0.19 ***	1.00								
10 Married	0.67	0.47	0.04 ***	0.00	-0.01	0.06 ***	0.34 ***	-0.06 ***	0.05 ***	0.25 ***	0.15 ***	1.00							
11 Previous Car Used	0.57	0.49	-0.11 ***	0.11 ***	-0.05 ***	-0.08 ***	-0.11 ***	-0.01	0.07 ***	0.01	-0.09 ***	0.04 ***	1.00						
12 Want Used	0.63	0.48	-0.18 ***	0.18 ***	-0.21 ***	-0.08 ***	-0.01	-0.06 ***	0.01	0.06 ***	0.05 ***	0.05 ***	0.13 ***	1.00					
13 Certified	0.33	0.47	0.17 ***	-0.12 ***	0.10 ***	0.13 ***	0.04 ***	0.06 ***	-0.01	-0.05 ***	0.08 ***	-0.02	-0.05 ***	-0.07***	1.00				
14 FEAT	0.46	0.44	0.12 ***	0.01	-0.13 ***	0.17 ***	0.23 ***	-0.03 **	-0.23 ***	0.08 ***	-0.25 ***	0.03 **	-0.04 ***	0.00	0.01	1.00			
15 RELB	0.35	0.40	0.11 ***	0.01	-0.10 ***	0.18 ***	0.22 ***	-0.03 **	-0.21 ***	0.06 ***	-0.20 ***	0.03 **	-0.07 ***	-0.05 ***	0.02	0.81 ***	1.00		
16 TRANS	0.19	0.29	0.10 ***	-0.00	-0.07 ***	0.17 ***	0.16 ***	-0.02	-0.15 ***	0.06 ***	-0.15 ***	0.03 **	-0.07 ***	-0.04 **	0.07 ***	0.64 ***	0.63 ***	1.00	
17 COMP	0.37	0.39	0.13 ***	-0.01	-0.09 ***	0.17 ***	0.23 ***	-0.04 **	-0.21 ***	0.08 ***	-0.20 ***	0.04 **	-0.07 ***	-0.05 ***	0.02 *	0.84 ***	0.77 ***	0.64 ***	1.00

Notes: \* p < .10, \*\* p < 0.05, \*\*\* p < 0.01; unpaired sample t-tests.

**Table 5.**  
**Model Estimation Results for Full sample using 3SLS**

	<b>M1a.</b> <b>PRICE</b>	<b>M1b.</b> <b>CERTI</b>	<b>M1c.</b> <b>FEAT</b>	<b>M1d.</b> <b>RELB</b>	<b>M1e.</b> <b>TRAN</b>	<b>M1f.</b> <b>COMP</b>
<b>A. VEHICLE CHARACTERISTICS</b>						
Miles	<b>-0.005</b> <b>(0.000)***</b>	<b>0.035</b> <b>(0.004)***</b>	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Model Year	<b>0.094</b> <b>(0.005)***</b>	<b>-0.625</b> <b>(0.078)***</b>	-0.009 (0.005)	-0.005 (0.006)	-0.001 (0.005)	-0.004 (0.006)
Price (Ln)		<b>6.475</b> <b>(0.705)***</b>				
<b>B. CERTIFICATION AND ONLINE INFORMATION</b>						
Certified	<b>0.147</b> <b>(0.019)***</b>					
Certified * Miles	0.000 (0.003)					
Certified * Luxury	0.009 (0.136)					
Features Information (FEAT)	0.071 (0.275)	<b>-3.800</b> <b>(0.893)***</b>				
Reliability Information (RELB)	<b>0.920</b> <b>(0.409)**</b>	<b>-8.715</b> <b>(1.125)***</b>				
Transaction Information (TRAN)	<b>-0.964</b> <b>(0.352)***</b>	<b>10.285</b> <b>(0.835)***</b>				
Comparatives Information (COMP)	<b>-0.468</b> <b>(0.282)*</b>	<b>7.903</b> <b>(1.296)***</b>				
FEAT * Certified	-0.023 (0.640)					
RELB * Certified	-0.023 (0.639)					
TRAN * Certified	0.038 (0.779)					
COMP * Certified	0.028 (0.391)					
<b>C. BUYER DEMOGRAPHICS AND PSYCHOGRAPHICS</b>						
Income	-0.001 (0.003)	-0.005 (0.013)	0.001 (0.002)	0.002 (0.002)	0.001 (0.002)	<b>0.004</b> <b>(0.002)**</b>
Minority	-0.005 (0.017)	0.221 (0.135)	<b>-0.060</b> <b>(0.017)***</b>	<b>-0.054</b> <b>(0.019)***</b>	<b>-0.041</b> <b>(0.016)**</b>	<b>-0.076</b> <b>(0.018)***</b>
Low education	<b>0.020</b> <b>(0.011)*</b>	<b>-0.195</b> <b>(0.089)**</b>	<b>-0.061</b> <b>(0.011)***</b>	<b>-0.055</b> <b>(0.012)***</b>	-0.015 (0.010)	<b>-0.047</b> <b>(0.012)***</b>
Gender (male)	0.000 (0.010)	-0.060 (0.075)	0.003 (0.010)	0.001 (0.011)	0.003 (0.009)	0.010 (0.010)
Age	-0.000 (0.000)	-0.002 (0.003)	<b>-0.002</b> <b>(0.000)***</b>	-0.001 (0.000)	-0.000 (0.000)	-0.001 (0.000)
Married	-0.020 (0.015)	0.120 (0.088)	-0.009 (0.011)	-0.009 (0.012)	0.002 (0.010)	-0.010 (0.012)
Previous Car Used	-0.012	<b>0.176</b>	-0.007	<b>-0.022</b>	<b>-0.018</b>	<b>-0.022</b>

*“The Impact of Online Information on the Purchase of Certified Used Cars”*

Want Used	(0.012) <b>-0.023</b> (0.012)*	<b>(0.074)**</b> <b>0.327</b> <b>(0.087)***</b>	(0.010) <b>0.021</b> <b>(0.010)**</b>	<b>(0.010)**</b> -0.016 (0.011)	<b>(0.009)**</b> -0.011 (0.009)	<b>(0.010)**</b> <b>-0.022</b> <b>(0.010)**</b>
<b>D. CONTROLS: OFFLINE INFORMATION SEARCH<sup>§</sup></b>						
Use of offline classifieds	-0.008 (0.005)	<b>0.076</b> <b>(0.033)**</b>	<b>0.061</b> <b>(0.004)***</b>	<b>0.038</b> <b>(0.004)***</b>	<b>0.025</b> <b>(0.004)***</b>	<b>0.037</b> <b>(0.004)***</b>
Use of offline personal source	<b>-0.016</b> <b>(0.004)***</b>	<b>0.120</b> <b>(0.028)***</b>	<b>-0.033</b> <b>(0.004)***</b>	<b>-0.023</b> <b>(0.004)***</b>	<b>-0.020</b> <b>(0.003)***</b>	<b>-0.023</b> <b>(0.004)***</b>
Constant	3.249 (0.079)***	-20.595 (2.398)***	-0.105 (0.078)	-0.087 (0.084)	-0.000 (0.071)	-0.211 (0.080)***
Observations	3823	3823	3823	3823	3823	3823
Fit statistics	R <sup>2</sup> = 0.60	R <sup>2</sup> = 0.13	R <sup>2</sup> = 0.62	R <sup>2</sup> = 0.47	R <sup>2</sup> = 0.28	R <sup>2</sup> = 0.49

**Notes:** <sup>§</sup> All models contain 125 car dummies for vehicle make-model-trim. Additional controls (not shown) for the price model M1a and choice of certification model M1b are market area (rural, small, metro), vehicle quality/condition (Satisfaction and NumProblems), purchase of additional warranty, and type of seller (new vehicle, used vehicle, other dealer). Other information controls include order of online search compared to dealer visits, and consumer psychographics.

All variables in panel B are modeled as endogenous and estimated using the instrumental variables technique with a surfeit of instruments.

\* p < .10, \*\* p < 0.05, \*\*\* p < 0.01; Standard errors shown in parentheses.

**Table 6.**  
**Model Estimation Results for Split sample using 3SLS**  
**DV: Price**

	Certified sample	Non-certified sample
<b>A. VEHICLE CHARACTERISTICS</b>		
Miles	<b>-0.004</b> (0.000) <sup>***</sup>	<b>-0.006</b> (0.000) <sup>***</sup>
Model Year	<b>0.090</b> (0.007) <sup>***</sup>	<b>0.097</b> (0.005) <sup>***</sup>
Price (Ln)		
<b>B. CERTIFICATION AND ONLINE INFORMATION</b>		
Features Information (FEAT)	-0.201 (0.132)	0.077 (0.132)
Reliability Information (RELB)	-0.081 (0.137)	<b>0.464</b> (0.171) <sup>***</sup>
Transaction Information (TRAN)	0.093 (0.142)	<b>-0.593</b> (0.163) <sup>***</sup>
Comparatives Information (COMP)	0.264 (0.209)	<b>-0.309</b> (0.159) <sup>*</sup>
<b>C. BUYER CHARACTERISTICS</b>		
Income	-0.000 (0.002)	<b>-0.003</b> (0.002) <sup>*</sup>
Minority	<b>0.048</b> (0.020) <sup>**</sup>	-0.003 (0.019)
Low education	-0.008 (0.015)	0.007 (0.012)
Gender (male)	<b>-0.039</b> (0.014) <sup>***</sup>	-0.007 (0.010)
Age	-0.000 (0.000)	0.000 (0.000)
Married	<b>-0.029</b> (0.014) <sup>**</sup>	-0.013 (0.012)
Previous Car Used	0.005 (0.011)	-0.009 (0.010)
Want Used	-0.014 (0.013)	<b>-0.029</b> (0.012) <sup>**</sup>
Observations	1226	2597
Fit statistics	R <sup>2</sup> = 0.85	R <sup>2</sup> = 0.72

**Notes:** § All models contain 125 car dummies for vehicle make-model-trim. Additional controls not shown include order of online search compared to dealer visits, use of offline classifieds and offline personal sources, market area (rural, small, metro), vehicle quality/condition (Satisfaction and NumProblems), purchase of additional warranty, and type of seller (new vehicle, used vehicle, other dealer).

All variables in panel B are modeled as endogenous and estimated using the instrumental variables technique with a surfeit of instruments.

\* p < .10, \*\* p < 0.05, \*\*\* p < 0.01; Standard errors shown in parentheses.

**APPENDIX**

**Table A1**  
**Results of the 3SLS estimation including the details of the controls used**

	<b>M1a.</b>	<b>M1b.</b>	<b>M1f.</b>	<b>M1c.</b>	<b>M1d.</b>	<b>M1e.</b>
	<b>PRICE</b>	<b>CERTI</b>	<b>FEAT</b>	<b>RELB</b>	<b>TRAN</b>	<b>COMP</b>
<b>A. VEHICLE CHARACTERISTICS</b>						
Miles	<b>-0.005</b> (0.000) <sup>***</sup>	<b>0.035</b> (0.004) <sup>***</sup>	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Model Year	<b>0.094</b> (0.005) <sup>***</sup>	<b>-0.625</b> (0.078) <sup>***</sup>	-0.009 (0.005)	-0.005 (0.006)	-0.001 (0.005)	-0.004 (0.006)
Price (Ln)		<b>6.475</b> (0.705) <sup>***</sup>				
<b>B. CERTIFICATION AND ONLINE INFORMATION</b>						
Certified	<b>0.147</b> (0.019) <sup>***</sup>					
Certified * Miles	0.000 (0.003)					
Certified * Luxury	0.009 (0.136)					
Reliability Information (RELB)	<b>0.920</b> (0.409) <sup>**</sup>	<b>-8.715</b> (1.125) <sup>***</sup>				
Transaction Information (TRAN)	<b>-0.964</b> (0.352) <sup>***</sup>	<b>10.285</b> (0.835) <sup>***</sup>				
Alternatives Information (ALTN)	<b>-0.468</b> (0.282) <sup>*</sup>	<b>7.903</b> (1.296) <sup>***</sup>				
Features Information (FEAT)	0.071 (0.275)	<b>-3.800</b> (0.893) <sup>***</sup>				
RELB * Certified	-0.023 (0.639)					
TRAN * Certified	0.038 (0.779)					
ALTN * Certified	0.028 (0.391)					
FEAT * Certified	-0.023 (0.640)					
<b>C. BUYER DEMOGRAPHICS AND PSYCHOGRAPHICS</b>						
Income	-0.001 (0.003)	-0.005 (0.013)	0.001 (0.002)	0.002 (0.002)	0.001 (0.002)	<b>0.004</b> (0.002) <sup>**</sup>
Minority	-0.005 (0.017)	0.221 (0.135)	<b>-0.060</b> (0.017) <sup>***</sup>	<b>-0.054</b> (0.019) <sup>***</sup>	<b>-0.041</b> (0.016) <sup>**</sup>	<b>-0.076</b> (0.018) <sup>***</sup>
Low education	<b>0.020</b> (0.011) <sup>*</sup>	<b>-0.195</b> (0.089) <sup>**</sup>	<b>-0.061</b> (0.011) <sup>***</sup>	<b>-0.055</b> (0.012) <sup>***</sup>	-0.015 (0.010)	<b>-0.047</b> (0.012) <sup>***</sup>
Gender (male)	0.000 (0.010)	-0.060 (0.075)	0.003 (0.010)	0.001 (0.011)	0.003 (0.009)	0.010 (0.010)
Age	-0.000 (0.000)	-0.002 (0.003)	<b>-0.002</b> (0.000) <sup>***</sup>	-0.001 (0.000)	-0.000 (0.000)	-0.001 (0.000)
Married	-0.020	0.120	-0.009	-0.009	0.002	-0.010

*“The Impact of Online Information on the Purchase of Certified Used Cars”*

Previous Car Used	(0.015) -0.012	(0.088) <b>0.176</b>	(0.011) -0.007	(0.012) <b>-0.022</b>	(0.010) <b>-0.018</b>	(0.012) <b>-0.022</b>
Want Used	(0.012) <b>-0.023</b>	<b>(0.074)**</b> <b>0.327</b>	(0.010) <b>0.021</b>	<b>(0.010)**</b> -0.016	<b>(0.009)**</b> -0.011	<b>(0.010)**</b> <b>-0.022</b>
Visit online before dealer	<b>(0.012)*</b> -0.003	<b>(0.087)***</b> -0.007	<b>(0.010)**</b> <b>0.428</b>	(0.011) <b>0.230</b>	(0.009) <b>0.436</b>	<b>(0.010)**</b> <b>0.554</b>
	(0.029)	(0.155)	<b>(0.011)***</b>	<b>(0.009)***</b>	<b>(0.011)***</b>	<b>(0.010)***</b>
<b>D. CONTROLS: OFFLINE INFORMATION SEARCH<sup>§</sup></b>						
Use of offline classifieds	-0.008 (0.005)	<b>0.076</b> <b>(0.033)**</b>	<b>0.061</b> <b>(0.004)***</b>	<b>0.038</b> <b>(0.004)***</b>	<b>0.025</b> <b>(0.004)***</b>	<b>0.037</b> <b>(0.004)***</b>
Use of offline personal sources	<b>-0.016</b> <b>(0.004)***</b>	<b>0.120</b> <b>(0.028)***</b>	<b>-0.033</b> <b>(0.004)***</b>	<b>-0.023</b> <b>(0.004)***</b>	<b>-0.020</b> <b>(0.003)***</b>	<b>-0.023</b> <b>(0.004)***</b>
<b>E. OTHER CONTROLS</b>						
Market_metro	0.019 (0.017)	<b>-0.120</b> <b>(0.071)*</b>				
Market_small	<b>0.032</b> <b>(0.012)***</b>	<b>-0.208</b> <b>(0.056)***</b>				
Overall vehicle satisfaction	-0.000 (0.001)	0.012 (0.016)				
NumProblems	0.002 (0.002)	-0.015 (0.012)				
Additional Warranty	0.004 (0.012)	-0.010 (0.049)				
New vehicle seller	<b>0.107</b> <b>(0.013)***</b>	<b>-0.689</b> <b>(0.114)***</b>				
Used vehicle seller	<b>0.075</b> <b>(0.014)***</b>	<b>-0.484</b> <b>(0.096)***</b>				
Constant	3.249 (0.079)***	-20.595 (2.398)***	-0.105 (0.078)	-0.087 (0.084)	-0.000 (0.071)	-0.211 (0.080)***
Observations	3823	3823	3823	3823	3823	3823
Fit statistics	R <sup>2</sup> = 0.60	R <sup>2</sup> =0.13	R <sup>2</sup> = 0.62	R <sup>2</sup> = 0.47	R <sup>2</sup> = 0.28	R <sup>2</sup> = 0.49

**Notes:** <sup>§</sup> All models contain 125 car dummies for vehicle make-model-trim. All variables in panel B are modeled as endogenous and estimated using the instrumental variables technique with a surfeit of instruments.

\* p < .10, \*\* p < 0.05, \*\*\* p < 0.01; Standard errors shown in parentheses.

## APPENDIX DETAILS OF THE INSTRUMENTAL VARIABLE REGRESSION

### **Endogeneity:**

In the presence of endogeneity, OLS estimates are biased and inconsistent (Maddala1992), whereas properly specified 2SLS/3SLS models are consistent. Further, in the presence of contemporaneous error-covariance, a method like 3SLS that makes use of the cross-equation correlations of the disturbances is asymptotically more efficient than 2SLS, and is the one we use.

The 3SLS procedure is as follows: (1) Each endogenous variable in the system (here, certification, certi X miles, certi X luxury, FEAT, RELB, TRAN, COMP, and Certi X Information) is regressed on all of the exogenous variables, and predicted values of the endogenous variables are calculated. (2) These predicted values are used as instrumental variables for the endogenous variables in OLS regressions, from which estimates of error terms and variance-covariance matrices are obtained. (3) Using the cross-equation disturbance correlation estimates, generalized least squares estimation is applied to a single large equation representing all the equations in the system.

### **Instruments:**

We identify a total of 14 instruments for the endogenous regressors in the structural model for price. We need at least as many instruments as there are endogenous regressors in each equation. However, while the asymptotic efficiency of the estimation improves as the number of instruments increases, but so does the finite-sample bias (Johnston and DiNardo 1997). Thus, from among the identified instruments, we choose the minimum subset needed to appropriately identify our system, using difference-in- Sargan tests, or the C statistic to eliminate any that are redundant by examining a subset of the orthogonality conditions.

Our set of instruments includes *Shop Certified* - a binary variable that describes that the buyer intentionally shopped only for a certified used vehicle; and interactions of *ShopCertified* with Miles and Luxury. Further, we use as instruments buyers beliefs regarding certification programs—there is value in certification, certification programs offer special financing/discounts, and previous experience with certified programs. We also include a dummy to indicate whether the car the buyer purchase just happened to be certified. We use an interaction between time spent on the Internet (hours/week) and *Shop Certified*. A set of four variables that measure the buyers' ratings of the importance of each type of information on a scale of 1 to 10, i.e. importance of FEAT, RELB, TRAN, and COMP, are used to instrument online information. Two variables are used to instrument for price- whether seller offered financing and whether buyer considered the vehicle purchased to be the best deal for the money.

### **Model comparison:**

We use the Hausman (1978) specification test, also known as the Wu-Hausman or Durbin-Wu-Hausman test to compare the OLS estimates with the IV estimates to determine if IV estimation is necessary. The null hypothesis is that the methods are equivalent, and the test statistic has a chi-square distribution with degrees of freedom equal to the number of potentially endogenous regressors. If the difference between the OLS and IV estimates is “large,” we conclude that OLS is not adequate. We also observe that the coefficients are significantly larger in the 3SLS model than the OLS model, indicating the need to appropriately account for endogeneity in our equations. Further, the DWH statistic  $X^2(11) = 21.03825$  ( $p=0.032$ ), is rejected.

**Validity and relevance of instruments:**

The validity of the overall set of instruments is assessed by examining the orthogonality (exogeneity) of the instruments to structural equation (price) using a Sargan test for overidentifying restrictions (Baum et al. 2003). We are unable to conduct this examination simultaneously for all the equations in our system, We therefore conduct two sets of 2SLS regression, one for the regression equation [1] and another for the choice of certification equation [2]. The null hypothesis is that, the excluded exogenous variables, the instruments, are uncorrelated to the regression error in the main equation. This statistic is chi-square distributed with degrees of freedom equal to the difference between the number of instruments and the number of endogenous variables. If the test statistic is significant and thus the null hypothesis is rejected, it indicates that either the model is misspecified or instrument orthogonality is violated; indicating that some of the instruments may be invalid. For the price equation, which contains 11 endogenous variables, we instrument using a set of 12 instruments, as described above. The Sargan statistic of identification/IV relevance =  $X^2(1) = 2.105$  ( $p=0.15$ ) is not rejected, supporting the null hypothesis that the price or outcome equation is well-identified. For the certification equation, which contains 5 endogenous variables (price and four online information measures), we instrument using 2 instruments for price and 4 for information. The Sargan statistic of identification/IV relevance =  $X^2(1) = 0.930$  ( $p=0.33$ ) is not rejected, supporting the null hypothesis that the certification or choice equation is also well-identified.

**Weak instruments:**

We also check for weakly identified solutions to ensure relevance of instruments. Weak identification occurs when the correlation between the instruments and the endogenous variable is small, and leads to problems of inference (Bound, Jaeger and Baker 1995). First, if the instruments and the endogenous variables are weakly correlated, then even a weak correlation between the instruments and the error in the original structural equation (which should be zero) can lead to large inconsistencies in the IV estimates; this is known as the “bias” issue related to weak instruments. Second, finite sample results can differ substantially from asymptotic theory. Specifically, IV estimates are generally biased in the same direction as OLS estimates, with the magnitude of this bias increasing as the  $R^2$  of the first-stage regression between the instruments and the endogenous variable approaches zero; this is known as the “size” issue related to weak instruments.

More recently, Stock and Yogo (2005) provide tests to determine if instruments are weak. However, their statistics are only computed for models with upto three endogenous. We therefore examine the following - Cragg-Donald test for model identification, the Anderson-Rubin test for the joint significance of the set of endogenous variables in our system of equations, the first stage F-stat, and the squared partial  $R^2$

If the first stage F-statistic (of the exclusion of the instrument set in the first-stage regression) exceeds the critical value (using 5% bias) from Stock and Yogo (2005), the instruments are deemed to be valid. For a single endogenous regressor, an F greater than 10 indicates sufficient relevancy (Staiger and Stock 1997). For a model with multiple endogenous regressors such as ours, the appropriate statistic to use is the the Cragg-Donald (1993) statistic, which equals the minimum eigenvalue of the matrix analog of the F-statistic from the first stage regression of IV estimation. The null hypothesis of this test is that the matrix of reduced form coefficients has rank= $K-1$ , where K is the number of regressors. If one finds support for the above null, then it would indicate that the system of equations is underidentified. In our model, we find strong evidence that the Cragg-Donald statistic ( $X^2(2) = 7.33$ ,  $p=0.03$  for price[1] and  $X^2(2) = 123.40$ ,  $p=0.00$  for certification

[2]) and the Anderson canonical correlation LR test statistic ( $X^2(2) = 7.32$ ,  $p=0.03$  for price [1] and  $X^2(2) = 121.45$ ,  $p=0.00$  for certification [2]) rejects the null, indicating that our system of equations is well-identified.

We report the partial  $R^2$  or the squared partial correlation between the excluded instruments and the endogenous regressors, taking correlations among the different instruments into account. Finally, the Anderson-Rubin (1949) test examines the joint significance of the multiple endogenous variables in our system of equations. The important point being that this test is robust even in the presence of weak instruments. This statistic with the null hypothesis that the endogenous regressors in the main structural equation are not significant, is rejected ( $F(12, 3665) = 2.36$ ,  $p=0.00$  for price [1] and  $F(6, 3673) = 8.50$ ,  $p=0.00$  for certification [2]).

### **Appendix References**

Anderson, T. and Rubin, H. (1949), “Estimation of parameters of a single equation in a complete system of stochastic equations,” *Annals of Mathematical Statistics*, 20, 46-63

Baum, C.F., Schaffer, M.E., Stillman, S., Wiggins, V. 2006. `overid`: Stata module to calculate tests of overidentifying restrictions after `ivreg`, `ivreg2`, `ivprobit`, `ivtobit`, `reg3`.  
<http://ideas.repec.org/c/boc/bocode/s396802.html>

Bound, John, David Jaeger, and Regina Baker (1995), “Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variables is Weak.” *Journal of American Statistical Association*, 90(430), 443–50.

Cragg, J.G. and S.G. Donald (1993), “Testing Identifiability and Specification in Instrumental Variable Models,” *Econometric Theory*, 9, 222 – 240.

Hausman, J. (1978), “Specification tests in econometrics”, *Econometrica*, 46(6), 1251-1271.

Johnston, Jack, and John DiNardo (1997). *Econometric Methods*, 4<sup>th</sup> edition. New York, NY: McGraw-Hill.

Maddala, G. S. (1992). *Introduction to Econometrics*. McMillan, New York, NY.

Staiger, Douglas and James H. Stock. 1997. “Instrumental Variables Regression with Weak Instruments.” *Econometrica* 65(3): 557–86.

Stock, James. H. and Motohiro Yogo (2005), “Testing for weak instruments in linear IV regressions,” In D.W.K. Andrews and J.H. Stock (eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge University Press, Cambridge.

**Table A2**

Instrument validity and relevance statistics for price equation [1]

Model	Certi	Certi Miles	Certi_Lu x	RELB	FEAT	TRAN	COMP	Certi X RELB	Certi X FEAT	Certi X TRAN	Certi X COMP
Partial R <sup>2</sup>	0.7155	0.1705	0.2660	0.2143	0.2337	0.0876	0.2273	0.0555	0.0472	0.0478	0.0396
First stage F (12, 3665)	767.93	62.79	110.66	83.30	93.16	29.31	89.85	17.93	15.13	15.34	12.61
P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Instrument validity and relevance statistics for certification equation [2]

Model	FEAT	RELB	TRAN	COMP	Price
Partial R <sup>2</sup>	0.23	0.21	0.08	0.22	0.04
First stage F (6, 3673)	185.71	159.39	50.08	174.64	22.52
P-value	0.00	0.00	0.00	0.00	0.00