RETURNS ON REPUTATION IN RETAIL E-COMMERCE

Jani Saastamoinen
University of Joensuu, Economics and Business Administration
P.O. Box 111, FI-80101, Joensuu, Finland
jani.saastamoinen@joensuu.fi

ABSTRACT

The industrial organization literature suggests that firms invest in reputation to earn price premiums. Data from online auctions has revealed that sellers are able to earn returns on their reputations. This paper examines online retail markets from the same perspective. We study, data from the markets of homogeneous consumer products listed in Pricegrabber.com in May 2008 is analyzed with a hierarchical regression model using OLS and quantile regression. Contrary to online auctions, the results indicate that in general, sellers do not earn returns on reputation in retail e-commerce. However, the evidence suggests that very large sellers and small sellers may benefit from their reputations in competition. Moreover, we discover that while an increase in the number of sellers lowers prices overall, the control groups are not affected by this but an increase in the number of small sellers lowers prices universally.

Keywords: retailing, e-commerce, competition, reputation effects, asymmetric information

1. Introduction

The Internet gives consumers unprecedented power in purchase decisions. Since the cost of search is minimal on the Internet, buyers can easily compare prices across several vendors before purchase. While price information is more accessible on the Internet, consumers face information-related risks in e-commerce. In e-commerce transactions, buyers disclose sensitive information, such as credit card details, to a seller. Furthermore, it is not possible to verify the quality of merchandise or the identity of a seller, because the merchandise is delivered after the seller has received a payment. Facing these problems of asymmetric information, buyers may need assurance that sellers do not cheat them. As a result, a good reputation or a widely recognized brand could be a valuable asset in e-commerce.

New online information services have reduced asymmetric information in retail e-commerce. To make price comparisons more convenient, several companies offer comparison shopping services. These websites enable comparison shopping on the Internet by providing up-to-date price quotes for various products. Very often these websites have reputation systems which collect and distribute information about the past activities of sellers. As a consequence, comparison shopping websites are highly competitive marketplaces where buyers are able to compare prices and risks associated with any particular seller. Comparison shopping websites are popular because many of them are among the 1000 most visited websites on the Internet[1]. Since large consumer flows can translate into higher revenues, firms have a solid financial incentive to participate in comparison shopping markets. In addition, they offer firms market information about the customer base and a low cost method of monitoring rivals.

Comparison shopping markets present a great opportunity to gain insights on the market structures of e-markets. The determinants of market structure are market concentration, product differentiation, the conditions of entry and exit and information [Jacobson and O'Callaghan-Andréosso 1996]. In comparison shopping markets, products are identical and the barriers to entry and exit are low. Therefore, market concentration and information will determine market structure. Since market structure determines pricing and profits, studying data from comparison shopping markets helps to understand how market concentration and information shape competition in e-markets.

Asymmetric information between buyers and sellers has inspired numerous researchers to inspect the relationship between a seller's reputation and prices in online auctions. Overall, these studies conclude that a good

[1] Examples of comparison shopping service websites include portals and search engines such as AOL (21), CNET (127), Google (2), MSN (5) and Yahoo! (1) and specialized comparison shopping websites such as Become.com (2786), Dealtime.com (465), Pricegrabber.com (870) and NexTag.com (546). A global web traffic rank in parentheses (retrieved July 1st 2008) as reported by Alexa (www.alexa.com), a company that tracks web traffic.

reputation allows some pricing power to sellers[2]. However, there is considerably less research on burgeoning retail e-markets. Baylis and Perloff [2002] study the e-markets of two homogeneous goods. They find that the firms that provide good service also set lower prices. They also use consumers' quality rankings from an outside source, but deem it "largely random information" that is worthless.

This study differs from the previous research because we examine the effect of reputation on seller's pricing in retail e-markets. We use cross-sectional data from over 6000 homogeneous goods markets obtained from Pricegrabber.com which is an online comparison shopping service[3]. The product categories in the sample include appliances, auto parts, children's products, cameras, computers, electronics, furniture, health and beauty products, indoor living products, musical instruments, outdoor living products, software, sporting goods, toys, TVs and video games. In addition to price data from these product categories, the study includes data from Pricegrabber.com's reputation system. The reputation data is generated by the buyers who elicit feedback on sellers. The reputation system reports this feedback as an aggregated rating score, a ratings history and verbal comments. We use quantitative measures to model a seller's reputation as a combination of the rating score and ratings history. We also consider the influence of auxiliary variables such as seller types and market concentration on a seller's pricing.

The rest of the paper is organized as follows. Section 2 provides the theoretical background and a brief literature review. In Section 3, we present data descriptions and research methodology. Section 4 presents the results from OLS and quantile regression. Section 5 discusses the results. Section 6 concludes the paper.

2. Literature Review

2.1. Foundations of Reputation and Trust in e-Commerce

Many markets suffer from detrimental effects of asymmetric information between trading agents [Akerlof 1970]. Even with free information, gathering and processing information is costly and prone to errors [Williamson 1973]. Another source of asymmetry is the contracts that determine the terms of trade. Agents cannot be certain that the counterparty fully obliges to the contract. To level these asymmetries, an agent's reputation is a signal of her trustworthiness. Cabral [2005] defines reputation as a situation, when a particular agent is expected to *be* something, whereas trust is defined as a situation, when a particular agent is expected to *do* something.

Game theorists formulated the theory of reputation building rationalize an incumbent firm's willingness to deter entry even if it is costly in the short-run [e.g. Milgrom and Roberts 1982]. A reputation arises from repeated action which other players in the game interpret as a commitment to continue to take the same action in the future [Dellarocas 2003]. In this setting, the incumbent firm develops a reputation of toughness by constantly deterring entry in order to discourage future entrants. Thus, the reputation based on the firm's past actions affects its future payoffs, and other players' probabilities about the firm's future actions are derived from its current reputation [Wilson 1985].

Building a reputation may provide tangible benefits for a firm. First, reputation can be viewed as an asset. In Klein and Leffler [1981] and Shapiro [1983], a firm invests in reputation by selling high quality products at loss initially (introductory pricing) but earns a price premium on the established reputation later. However, to be qualified as an asset implies that the established reputations can be bought. For this reason, Mailath and Samuelson [2001] argue that a reputation may not be a good signal of quality because incompetent firms can buy good reputations. Second, Klein and Leffler [1981] suggest that consumers view reputation as protection for contractual obligation. Returns on reputation induce a firm to maintain good quality because the profit stream from good quality products exceeds the gains from cheating. Hörner [2002] argues that this does not provide sufficient incentives to maintain good quality. Instead, competition provides such incentives by creating an outside option to buyers who can patronize the firm's rival, if they detect cheating.

The amount of trust a buyer places on a seller depends on a seller's reputation [Resnick et al. 2000]. In other words, the buyer's beliefs on the risks involved in a transaction with the seller are based on the assessment of the seller's reputation. This can be based on a personal transaction history with the seller. Another source could be learning from other agents [Dua et al. 2009]. The incentive structure of the game is influenced by a threat of retaliation or reciprocity. For example, if a seller cheats a buyer (or vice versa), the seller may lose all future transactions with the buyer, or there will be legal consequences from a faulty action. In addition, the existence of reputations may discourage the market entry of less reliable sellers (or buyers) [Resnick et al. 2006].

Although the lower information costs of e-markets could make market incumbents better informed about the market in many respects, the nature of e-commerce raises concerns about the trustworthiness of a trading partner. In e-commerce, buyers and sellers conduct business through a website (or other electronic channel). As a result, lack of

---

[2] See Sun (2008) for a concise review of results.

[3] See www.Pricegrabber.com.

direct contact between a buyer and seller in online transactions leads to uncertainty about the identity of the trading partner and product quality [Ba and Pavlou, 2002]. More precisely, two main concerns are the loss of money and privacy [Resnick et al. 2000]. Apart from information goods, a buyer cannot inspect the merchandise before purchase. Showing pictures of the merchandise and suitable descriptions reduce some informational problems, but does not resolve all quality concerns [Bland et al. 2007]. Moreover, a seller's online store may not give any information about the seller's quality, so it is harder to verify quality in e-markets than in conventional markets.

Different product types increase the risks of asymmetric information in online transactions. Kotler and Keller [2006] identify four product types with unique characteristics. First, convenience goods, such as personal hygiene products, are relatively inexpensive. They are bought regularly, so the quality is usually known in advance. Hence, it is plausible that these are bought in bulk quantities online [Thirumalai and Sinha 2009]. Second, shopping goods, such as refrigerators, are more expensive than convenience goods and thus, they require price and quality comparisons before purchase. Third, specialty goods which are branded goods or they have a very specific usage warrant a price premium over regular shopping goods. For example, Apple's Mac computers are specialty goods. Due to increasing price and specificity, the risks associated with the product type increase as we move from convenience goods to specialty goods [Thirumalai and Sinha 2009]. The fourth product category, unsought goods, requires active selling efforts because a buyer does not recognize the need for an unsought good. As a result, e-commerce may not be the optimal sales channel for unsought goods.

More concerns surface in payment of a purchase. In general, buyers prefer that the information that a seller obtains from them will be kept private and its use in marketing purposes limited [Brown and Muchira 2004]. Most retail e-commerce transactions are paid by credit card. Disclosing sensitive information, such as a credit card number, involves the risks of misconduct. The ease of switching one's identity on the Internet accentuates these problems. The shipping of goods is also problematic because a buyer can only trust that a seller obeys the contract they have entered into. This does not mean, however, that only the seller's trustworthiness is a concern. Third parties, such as criminals or marketing companies, may gain access to sensitive information by illicit means. The delivery agency may also lose the ordered purchase which strains the buyer-seller relationship.

As trust in transactions is built on an agent's reputation, the locality of transactions is an important contributor to reputation building in conventional markets [Resnick and Zeckhauser 2002]. When transactions take place in the same physical environment, contacts with the seller are frequent. Thus, the seller's identity is known and buyers are able to inspect the merchandise before purchasing it. They are also able to learn from each other's experiences with the seller, so the word-of-mouth contributes to the seller's reputation. Moreover, the risks of privacy and shipping are negligible because buyers monitor the payment and organize the transportation of purchased items by themselves. Sellers may signal reputation by acquiring retail space in upscale locations. In addition to location, sellers in conventional markets can borrow reputations (certifications), buy reputations (an acquisition of an existing brand) or leverage existing reputations to new markets.

Compared with conventional markets, location advantages are less important in reputation building e-markets because firms lack the physical retail space. Nevertheless, some parallels in e-markets exist because several websites attract large consumer flows daily. First, comparison shopping websites such as Bizrate.com or Pricegrabber.com.com provide marketplace platforms for e-retailers. Second, on-line auction sites, most notably eBay, are global centers of consumer-to-consumer e-commerce as well as a sales channel for small-scale e-retailers. Third, web portals such as MSN, Yahoo! or CNET offer a wide range of services to their customers and are therefore potential locations for e-commerce. Furthermore, some websites for specific interest groups also support platforms for e-commerce. For example, Discogs.com, an electronic database for discographies, offers a marketplace function for its registered members. Finally, well-known e-commerce vendors, such as Amazon.com or Play.com, have set-up marketplaces where affiliate sellers can benefit from the brand and customer base of the marketplace provider.

Due to the risks involved in an e-commerce transaction, consumers need to be assured that they can trust the seller. Fast shipping, the traceability of purchased good, generous guarantees and return policies can foster trust. Alternative ways to pay for the product could also signal trust because credit cards and some e-payment solutions such as PayPal provide consumer protection in purchases[4]. Approval ratings from an impartial third party are another way to signal trust. As in financial markets, where the ratings from companies such as Standard & Poor's serve this purpose, e-commerce merchants use certifications from the providers of Internet security, such as McAfee or VeriSign, to communicate that measures have been taken to protect consumers in online transactions [Odom et al. 2002]. However, these will only be understood by the consumers who are familiar with the Internet security certifications [Kim and Benbasat 2003]. Other ways to signal trust are not as easily verifiable. Easy navigation and a

[4]The seller is also protected because the creditor bears the risks of consumer insolvency.

“professional look” of a merchant’s website could be crucial factors in purchase decisions[5]. Customer service, the ease of contacting the seller, effective communication, and consumer empowerment are also important in the reputation building process. Furthermore, introductory prices can be used to acquire frequent transactions that are instrumental in reputation building among consumers [Borenstein and Saloner 2001].

Since building a reputation is dynamic, costly and time-consuming effort, there are alternative shortcuts to establish this goal. Established businesses in conventional markets often leverage their offline reputations in on-line markets by expanding the offline business model to e-markets. Another strategy is to buy a reputation by an acquisition of an established online business or its brand or by a franchising agreement. Small enterprises may find it profitable to sell their merchandise under the umbrellas of strong, established on-line brands that offer some protection to buyers in the marketplace purchases.

A market maker can level the information asymmetry by collecting and distributing performance histories. Performance histories, which are user accounts on interactions between buyers and sellers, are an integral part of feedback mechanisms. A consistent performance history could provide the same proof as repeat purchases for quality-conscious buyers on a seller’s commitment to maintain high quality service [Rao and Bergen, 1992]. They provide quantitative (e.g. a length of history) and qualitative (e.g. a description about performance) information about transactions. Performance histories can be used to evaluate potential risks involved in a transaction [Resnick et al. 2000]. They can also be viewed as a signal of a seller’s capacity [Lin et al. 2006].

2.2. Reputation Systems

To address the problems of asymmetric information, e-commerce marketplaces have devised reputation systems that provide information about market incumbents past actions[6]. One can think of this as the digital word-of-mouth [Dellarocas 2003]. Resnick et al. [2000] define a reputation system as a system that “collects, distributes and aggregates feedback about participants past behavior”. To be effective, a reputation system should be long-lived and efficient in distributing information about reputations. Such a system alleviates asymmetric information problems between trading partners and encourages behavior that fosters trust. Usually, a reputation system reports a seller’s or buyer’s numerical rating score which may be accompanied by text comments and a ratings history. These assessments are provided by the sellers or buyers that have transacted with the seller or buyer.

Participation in a reputation system could signal that an agent is a trustworthy trading partner. Zhou et al. [2008] present a model for online markets where they show that a reputation system can reduce asymmetric information in an online market and replicate the results of Shapiro [1983]. An efficient reputation system provides incentives to fulfill contractual obligations. Also, it must create incentives to participate and report truthful feedback through the reputation system. Bakos and Dellarocas [2003] show that an online reputation system can be more efficient in enforcing desired behavior than a threat of litigation. On the other hand, building a reputation in one marketplace creates switching costs for established sellers because reputations are not transferable between competing marketplaces [Melnik and Alm 2002]. For this reason, a marketplace operator has an incentive to encourage participation in the marketplace’s reputation system because it creates a lock-in for sellers (and buyers, if they also act as sellers and vice versa).

Despite their benefits, reputation systems are not a foolproof solution to the problems of asymmetric information. While feedback mechanisms rely mostly on quantitative measurements, awarding feedback is subjective [Pavlou and Dimoka 2006]. As a result, a sale of a homogeneous product can receive very different reviews because agents have heterogeneous preferences. By the law of large numbers, feedback converges to some value but the problem exists with small amounts of feedback. Moreover, herding among feedback givers could bias reputations. Herding occurs when the private signals of agents become correlated with the public signals that they observe [Banerjee 1992]. If herding occurs, a buyer may let the public opinion published by the reputation system influence her assessment. However, a study by Cabral and Hortaçsu [2006] does not support this argument.

Another problem emerges from voluntary participation, because eliciting feedback imposes an incremental cost to a transaction. This compares with contributing to a public good: avoiding the cost of giving feedback creates an incentive to free-ride on the information that other agents provide. Feedback may also become biased because only the extraordinarily bad or good performances are reported [Pavlou and Dimoka 2006]. Moreover, the fear of retaliation could deter eliciting negative feedback.

Even more damaging to reputation systems could be the proliferation of markets for feedback. Brown and Morgan [2006] describe situations where eBay’s feedback mechanism is manipulated by selling merchandise that is essentially worthless in exchange for positive feedback. As a result, the accumulated positive feedback can be used

[5] This may not help because websites that have a “professional look” are easy to forge (Kumaraguru et al. 2006).
[6] Williamson (1973) suggests that “a simple performance record can be maintained and a priori probabilities successively revised” to reduce asymmetric information about “the true characteristics of economic agents.”

to signal a good reputation in fraudulent listings of valuable items. Another way to go around reputation systems is *shilling* which is usually explicitly forbidden in online marketplaces. For example, a seller could act as a buyer and purchase a product from her own online store and return positive feedback for herself. This could undermine the value of the reputation systems for buyers because the inability to manipulate one's own reputation is partly responsible for the value of reputation [Standifird 2001]. As reputation systems do not distinguish between the monetary values of sold items, it is possible that a seller amasses a good reputation by selling inexpensive items but eventually cheats at a sale of a valuable item [Livingston 2005]. Distinguishing manipulation from regular business practices may be difficult though. For example, a seller offering a used CD for a nominal fee of 1 cent may be clearing inventory, or applying the introductory pricing strategy.

A major handicap for reputation building on the Internet is that changing one's identity is relatively costless. For example, creating a new seller identity in online auctions requires only registration. In the online retail industry, comparison shopping services, such as Pricegrabber.com.com or Yahoo!, offer packages that enable a quick set-up of an online store at low costs. While these features guarantee low entry costs, asymmetric information between buyers and sellers may deter frequent switching of one's online identity in retail markets. Switching an identity means that a seller must start building its reputation again from the beginning because the accumulated ratings under the previous identity are not transferable. For this reason, the investment costs of reputation building serve as an entry costs to another market. Online marketplaces also create switching costs for the market incumbents by not allowing a transfer of reputations between marketplaces [Brown and Morgan 2006]. Therefore, marketplace operators have an incentive to encourage sellers into reputation building, because they gain more sales fees from the locked-in sellers.

The issues of trust and reputation can lead to the problems of adverse selection and moral hazard in e-markets. A strong positive reputation can be viewed as an insurance against opportunistic behavior [Standifird 2001]. The cost of adverse selection can be that sellers receive lower prices for their goods, or even unraveling of the markets in the extreme cases [Akerlof 1970; Dewan and Hsu 2004]. As the full price of a product is a combination of a purchase price, search costs and costs of a disappointing purchase, a good reputation can mitigate the costs of a disappointing purchase [Kim and Xu 2007]. Thus, a reputable seller could enjoy a price premium over its less trustworthy rivals. Melnik and Alm [2002] suggest that reputation can raise barriers to entry in an e-market because new entrants may find it impossible to compete with the established reputable sellers. Indeed, a study of eBay auctions by Lin et al. [2006] suggests that the population of sellers with high reputation scores has higher growth rate than the sellers with lower reputation scores. Interestingly, Professional eBay Sellers Alliance, a trade association of high-ranking eBay sellers, complains that exactly the opposite is taking place in eBay. They claim that the marketplace does not provide enough incentives for sellers to invest in the measures that improve seller reputations[7].

2.3. Empirical Evidence

Popular online auctions have become a widely-used data source for researchers because reputation systems are commonplace in auction sites[8]. The empirical evidence indicates that the returns on reputation in online auctions vary considerably. Standifird [2001] finds only limited evidence of price premiums for a seller with a good reputation, but a highly negative reputation forces a seller to sell items at discount. Furthermore, he finds evidence that a negative reputation has more impact on a buyer's purchase decision than a positive reputation. This finding is supported by Ba and Pavlou [2002]. They find little evidence of a positive correlation between rating scores and price premiums, but a statistically significant impact of a negative rating on a seller's price exists when the auctioned items are expensive.

In contrast to these findings, Melnik and Alm [2002] show that a seller's reputation has a small positive impact on prices in the auctions for gold coins. They argue that the price premium from a reputation is likely to grow along the value of an auctioned object. Dewan and Hsu [2004] also report similar findings in the auctions for collectible stamps. They estimate that quality uncertainty lowers the prices of auctioned stamps by 10-15%. They interpret this as the evidence of effective dealing with Akerlof's [1970] "lemons problem". Using coin auctions, Lucking-Reiley et al. [2006] discover that positive ratings give a mild boost to the final price whereas negative comments depress final prices. Moreover, Cabral and Hortaçsu [2006] report a 5% price premium for a better reputation score using coin, laptop computer and teddy bear auctions, while Livingston [2005] finds decreasingly increasing returns on seller's positive reputation in eBay auctions. Using quantile regression, Sun and Hsu [2007] detect a nonlinear response to a seller's reputation as buyers place more emphasis on the reputation when bid values are high. Since buyers can also be sellers in eBay auctions, Zhang [2006] investigates the impact of a reputation accumulated either

[7] See Professional eBay Sellers Association (2007): "Unhealthy Marketplace Dynamics – Seller Perspective".
[8] See Dellarocas (2003) for a summary of this literature.

as a buyer or a seller on auction prices. The results show that a reputation accumulated as a buyer does not matter as a seller, but a reputation as a seller matters. Furthermore, negative feedback has a greater impact on the final price than positive feedback.

Attempts to control for quality uncertainty have produced muted results on a seller's ability to earn returns on reputation. McDonald and Slawson [2002] inspect auctions for new or mint-condition Harley-Davidson Barbie-dolls. They find that a seller with a better reputation attracts more bids and commands higher prices. Studying auctions for Pentium processors, Houser and Wooders [2006] detect only a small 0.17% increase in the final price when a seller's reputation score increases by 10%[9]. In contrast, negative feedback lowers the final price by 0.24%. Inspecting auctions for gift-cards which have a pre-determined value, Pate [2006] discovers that a seller's reputation accounts for only 1% of the sales price. Moreover, her evidence indicates that the positive returns on reputation do not increase significantly with the monetary value of a gift card. Resnick et al. [2006] conduct an experiment in which an established seller of vintage postcards, who has a strong reputation in eBay, sets up new seller identities and sells vintage postcards under the established identity and the new identities. They find that the established seller commands a price premium of 8.1% over the new sellers. Using data from the auctions for new Apple iPod mp3-players, Sun [2008] reports 0.1% increase in the final price in response to 10% increase in the reputation score.

The empirical research on retail e-markets has been interested in price dispersion on the Internet because online markets were expected to be highly competitive. The existence of online price dispersion is well-documented in the literature[10]. The most common explanation for price dispersion is the cost of obtaining information [e.g. Stigler 1961; Varian 1980]. Since online search costs are low, other explanations have emerged. Smith & Brynjolfsson [2001] suggest that seller brands may explain price dispersion by allowing branded sellers to charge higher prices than other sellers. On the other hand, Ancarani and Shankar [2004] argue that seller characteristics determine price dispersion. That is, traditional, multi-channel and Internet retailers, charge different prices for the same goods. Should differentiation be a successful strategy, buyers must place some value on the services these seller types provide. As e-markets mature, differential pricing across seller types may not be possible. For example, Gan et al. [2007] find that price levels between pure-play Internet and multi-channel retailers do not differ in the Singaporean online groceries market. Pan et al. [2002] use hedonic regression to measure the impact of a seller's service quality on prices of various goods. In general, shopping convenience, reliability and shipping and handling are positively related to prices, whereas the amount of offered product information has a negative impact on prices. They conclude that sellers are not always able to charge higher prices by offering quality service. However, they point to the possibility that online trust or branding could explain price premiums in e-commerce.

Unlike online auctions, the extant research on the influence of a seller's reputation in retail e-markets is very thin. In a pioneering study, Baylis and Perloff [2002] observe price developments in two homogeneous consumer electronics products. Their findings are startling. They observe that "good firms" charge lower prices, while "bad firms" charge higher prices. Moreover, the relative price positions among firms do not change over time, which implies that periodic sales do not take place. Using a survey data from the customers of an online bookstore, Kim and Xu [2007] find that a seller's reputation can reduce a buyer's price sensitivity. Another survey by Fuller et al. [2007] suggests that seller ratings, which are provided by a reputation system, do not have a lasting impact on a buyer's decision making. In fact, buyers place more emphasis on direct personal experience either from the previous transactions with the seller or the information a buyer receives from the seller's website. These findings imply that the empirical results from online auctions may not be directly applicable to retail e-markets.

As the cost of obtaining information about a seller (or at least the information that is available on the Internet) is very low on the Internet, increasing the number of market incumbents should lower prices. Two factors contribute to this. First, a greater number of sellers in the market increases supply which lowers the price provided that the demand does not change. Second, a greater number of sellers in the market increases the likelihood of better fit for buyers. As a consequence, buyers are able to select a combination of attributes, such as a price-quality combination, that has the best fit to a buyer's preferences. The empirical evidence from eBay auctions supports this [McDonald and Slawson 2002]. In retail markets, Baye et al. [2004] and Leiter and Warin [2007] find that price dispersion measured by the price gap between the two lowest prices decreases when the number of sellers in the market increases.

---

[9] According to the authors, mostly unused processors were auctioned, which alleviates the perceived uncertainty over the product quality.

[10] See Pan et al. 2003 for a review of earlier studies.

2.4. Research Hypotheses

Based on the literature review, we propose the following research hypotheses[11]. First, the theory and evidence from online auctions predict that sellers earn returns on reputation. Hence, we examine whether or not sellers in online retail markets are also able to earn price premiums with their reputations. In addition, switching identities is easy on the Internet. For this reason, the length of a seller's ratings history may signal the seller's commitment to stay in the market. Thus,

Hypothesis 1A (H1A). Better reputations measured by rating scores enable price premiums.

Hypothesis 1B (H1B). Longer market histories measured by ratings histories enable price premiums.

Second, the literature suggests that branded sellers or well-known sellers may earn price premiums in online markets. Moreover, since brands may signal a seller's trustworthiness to buyers, relatively unknown sellers (storefronts) may compete with low prices to gain sales[12]. For this reason,

Hypothesis 2A (H2A). Well-known sellers charge higher prices than other sellers.

Hypothesis 2B (H2B). Storefronts charge lower prices than other sellers.

Third, the economic theory predicts that price competition intensifies as the number of market incumbents increases. As a continuation to H2B, we hypothesize that an increase in the relative number of storefronts in a market depresses prices. As a consequence,

Hypothesis 3A (H3A). Prices decrease as the number of market incumbents increases.

Hypothesis 3B (H3B). Prices decrease as the ratio of storefronts to the total number of market incumbents increases.

Finally, the importance of a seller's reputation in a consumer's purchase decision may vary across product types. We consider three broad product categories, convenience goods, shopping goods and specialty goods. Convenience goods are less risky in terms of price and quality than shopping and specialty goods. For this reason,

Hypothesis 4. Returns on reputation, if they exist, are greater to the sellers of shopping and specialty goods than to the sellers of convenience goods.

We test the validity of these research hypotheses by analyzing data from retail e-markets. Data is obtained from Pricegrabber.com.com which is a popular comparison shopping website. In the next section, we describe the research methodology and data in more detail.

3. Data Description and Analysis

3.1. Methodology and Variables

In this study, the proposed hypotheses are being tested on a data set that was collected from Pricegrabber.com. The data set is a cross-sectional random sample of posted prices and reputation scores from various product categories. The product categories include appliances, auto parts, children's products, cameras, computers, electronics, furniture, health and beauty products, indoor living products, musical instruments, outdoor living products, software, sporting goods, toys, TVs and video games. All products are new to avoid product heterogeneity. The data set is a random sample of the products that were available in Pricegrabber.com.com in May 2008. By using a computer program designed for the purpose, the data set was collected from Pricegrabber.com. Data for each product category was extracted within the same day, but the entire process took several days to accomplish.

Altogether, the sample consists of 6885 different markets for homogeneous goods. Data contains information on prices, the number of sellers in a market, seller types, seller rating scores and ratings histories for each market. From this data, we construct variables for regression analysis. The price (*PRICE*) is the dependent variable in regression analysis. A seller's rating score (*RATE*) and a seller's ratings history (*HIST*) are reputation variables while market thickness (*THICK*) and the storefront ratio (*SFR*) are market variables. Descriptive statistics together with variable descriptions are shown in Table 1. The price data consists of 18044 posted price quotes in U.S. dollars. All products are new and homogeneous which eliminates the possibility of product differentiation as a source of price differences.

The rating score is a measure for a seller's reputation. *RATE* is a decimal value which is provided by consumers who review the seller's overall performance in a transaction in 1 (the lowest) to 5 (the highest) scale after the transaction has concluded. As a result, the rating score is an aggregate value of the overall consumer opinion on the particular seller. Although rating scores provide an easily quantifiable measure for seller reputations, they are not without limitations. First, a buyer's review on a seller's performance is subjective. Due to buyer heterogeneity, the same level of service performance may lead to different ratings. Second, rating scores are often accompanied by verbal comments. These may contain very important information about a seller's conduct which is not captured by the numerical score. For example, a verbal comment on charging the credit card without shipping the order or a

[11] The proposed hypotheses are made with the standard ceteris paribus assumption.

[12] Seller types will be defined later.

significant delay in delivery could lead to a similar rating, but send a starkly different signal to other buyers. Pavlou and Dimota [2006] show that accompanying verbal comments have a limited effect on prices in eBay auctions, but quantitative measures of a seller's reputation are more important. Third, leaving feedback is optional, which may lead to biased feedback. This could be a result if only exceptionally good or really bad performances are reported. Finally, sellers may manipulate their rating scores as suggested by Brown and Morgan [2006]. However, this is likely to be a problem that is specific to online auctions, because e-retailing is characterized by large volumes of new merchandise, whereas auctioned items are often used and sales volumes are low.

Table 1: Descriptive statistics

Variable	*Description*	*Mean*	*Median*	*Std. Dev.*	*Min.*	*Max.*
RATE ***(RATE1)***	A seller's average all-time rating.	4.01 (3.99)	4.35 (4.37)	0.87 (0.95)	1.00 (1.00)	5.00 (5.00)
HIST* ***(HIST1*)***	The number of ratings a seller has received.	2171.76 (949.40)	373 (133)	5004.39 (2784.84)	10 (10)	36226 (36226)
PRICE ***(PRICE1)***	The list price of an item in USD.	408.67 (342.68)	112.05 (64.95)	1023.13 (1140.30)	0.01 (0.20)	41049.84 (40209.95)
SFR ***(SFR1)***	The ratio of storefronts to all market incumbents.	0.15 (0.13)	0.07 (0)	0.20 (0.20)	0 (0)	1 (1)
THICK ***(THICK1)***	The number of market incumbents in a market.	10.24 (8.27)	7.00 (4)	9.45 (9.20)	2 (1)	49 (49)
*Histories below 10 are not reported. The figures in parentheses refer to the markets where only one seller was active at the time the observations were made.						

Another way to estimate a seller's trustworthiness is to consider ratings histories. *HIST* provides an indicator of how long a seller has been active under the same identity. The longer the ratings history, the more reliable the seller is. Descriptive statistics indicate that ratings histories are heavily skewed to the left. The figures suggest that most sellers are relatively new to the market, or their sales volumes in Pricegrabber.com are low. Cabral and Hortaçsu [2006) argue that a frequency of feedback is a good approximation for a frequency of sales in eBay auctions. Hence, a ratings history may also be interpreted as a measure of sales in online retail markets.

Only one active seller could be found in a large number of markets at the time the observations were made. Since these markets cannot be used in the analysis, they were dropped from the sample. T-tests and Mann-Whitney-tests are conducted to determine whether data becomes biased in the process. The omitted portion of data is denoted by *HIST1*, *RATE1* and *PRICE1* variables. The test results are reported in Table 2. The first test between the rating scores (*RATE*, *RATE1*) is inconclusive. However, the summary statistics in Table 1 indicate that the data for both groups resembles each other. The second test between the ratings histories (*HIST*, *HIST1*) leads to the rejection of the equality assumption. The summary statistics show that the average ratings histories are longer in more competitive markets. The third test between the prices (*PRICE*, *PRICE1*) suggests that the equality assumption should be rejected. In this case, the markets with more active sellers display higher prices on average than the markets of one active seller. As a conclusion of this analysis, the omission biases data. However, we do not believe this to be a major cause for concern because the rating score, which is the most important variable in our analysis, appears to be roughly unaffected. Moreover, since average prices are higher in the remaining sample, this may help detect price premiums because a buyer's risk aversion presumably increases with the price.

Table 2: Results for T-tests and Mann-Whitney Tests

Hypothesis	*T-Test*	*Mann-Whitney*	*Conclusion*
RATE = RATE1	2.215**	1.166	Rejected/Accepted
HIST = HIST1	10.044***	20.298***	Rejected/Rejected
PRICE = PRICE1	3.456***	14.789***	Rejected / Rejected
*** p-value<0.01; ** p-value<0.05; * p-value<0.1.			

Market thickness measures the number of sellers in a market. Market thickness increases when the number of sellers in the market increases. In the remaining data, market thickness varies from 2 to 49 with the mean at 8.27, the median at 7 and the standard deviation of 9.45. These figures are considerably less than the average of 17.5 sellers in a market reported by Leiter and Warin [2007], who also obtain their data from Pricegrabber.com. A likely reason for the discrepancy is the sampling method, because their sample consists of the most popular products listed in Pricegrabber.com. Furthermore, dividing the number of storefronts (their definition is explained later) in a market by the market thickness of the respective market, we obtain a variable for the storefront ratio (*SFR*]. On average, 7% (median] to 15% (mean] of market incumbents are storefronts.

To facilitate a direct comparability between markets for different products, we use standardized variables. The downside to this procedure is that it is impossible to give straightforward interpretations for estimated coefficients although the direction and magnitude of influence can be given. In general, a standardized variable $z_{i,j}$ is obtained by

$$z_{i,j} = \frac{x_{i,j} - \bar{x}_j}{STD(x_{i,j})}, \tag{1}$$

in which $x_{i,j}$ is an observation of the variable x_i in the market *j*, $\bar{x}_j$ is the mean of *x* in the market *j* and $STD(x_{i,j})$ is the standard deviation of *x* in the market *j*. Standardization concentrates the observations around zero. We denote the standardized variables with the letter *z* in front of a variable. The standardized variables include the price (*zPRICE*], rating score (*zRATE*], ratings history (*zHIST*], market thickness (*zTHICK*] and storefront ratio (*zSFR*]. Table 3 lays out descriptive statistics for the standardized variables.

Table 3: Descriptive Statistics for Standardized Variables

Variable	*Mean*	*Median*	*Std. Dev.*	*Min.*	*Max.*
zRATE	0.000	0.285	0.883	-4.193	1.911
zHIST	0.000	-0.395	0.903	-1.562	4.513
zPRICE	0.000	-0.165	0.909	-2.888	5.327
zSFR	0.000	-0.410	1.000	-0.769	4.260
zTHICK	0.000	-0.343	1.000	-0.872	4.102

3.2. Seller Types

Firms of all sizes compete in comparison shopping markets. Pricegrabber.com has two fundamentally different seller types which are based on a seller's a choice over a sales channel. Pricegrabber.com defines "merchants" the firms that run their own websites. Merchants register to Pricegrabber.com to lure in customers but process commercial transactions through their own e-commerce systems. Pricegrabber.com defines "storefronts" the firms that do not run their own websites. Instead, the comparison shopping website processes commercial transactions between consumers and storefronts. Pricegrabber.com sets its fees in a way that merchants pay a fee for each click-through whether or not this leads to a purchase, but storefronts pay fees only for purchases. The click-through fee for merchants is significantly lower than the purchase fee for storefronts[13]. As a result, large volume sellers benefit from being merchants, whereas low volume sellers are induced to select the storefront package.

We devise dummy variables to distinguish between the different seller types. The dummy variable *SF* takes the value 1 if a seller is a storefront and the value 0 if a seller is a merchant. Further distinctions are based on the level of sales. We use the annual list (2007] of *The Internet Retailer* to separate large, branded sellers from the rest of the seller population[14]. The dummy variable *TOP500* takes value 1 if a seller is among the 500 hundred largest retail e-commerce sellers measured by the value of their annual sales in the United States. Thus, the control groups are at the opposite sides of the seller spectrum: storefronts are small players, whereas Top500-sellers are household names with wide brand recognition among consumers.

[13] See http://www.Pricegrabber.com.com/sell_here.php .
[14] See http://www.internetretailer.com/Top500/list.asp .

3.3. Product Type

The data set contains a diverse selection of consumer goods. To carry out robustness analysis, we construct a subset of data by assigning selected products in three categories as described in Kotler and Keller [2006][15]. Convenience goods are the lowest risk items for consumers, whereas the risk level increases when consumers buy shopping and specialty goods [Thirumalai and Sinha 2009]. The convenience goods data set (*CONV*) includes baby bottle feeding formula, deodorant, cleaning products, printer paper and storage media. The shopping goods data set (*SHOP*) includes desk-collections, flat-panel LCD-monitors, mp3-players, plasma/LCD-televisions and refrigerators. The specialty goods data set (*SPCL*) includes binoculars, electric guitars, microphones, network surveillance equipment and SLR-lenses.

Descriptive statistics for these product types are shown in Table 4. Price statistics verify that the average price for convenience goods is significantly lower than the average prices for the other two product categories. Ratings statistics display less severe absolute differences in averages but the distribution of the rating scores is more concentrated for specialty goods than for the other two product categories. History statistics indicate that the sellers of shopping and specialty goods have been active in their respective markets for a longer period of time than the sellers of convenience goods. The market variables verify that the average number of sellers is lower in the specialty goods markets than either the convenience or shopping goods markets reflecting the specific nature of specialty goods. In addition, the low median of the storefront ratio suggests that specialty goods are not a niche strategy that storefronts favor.

Table 4: Descriptive statistics for Product Categories

Variable		*PRICE*			*RATE*			*HIST*	
Category	*CONV*	*SHOP*	*SPCL*	*CONV*	*SHOP*	*SPCL*	*CONV*	*SHOP*	*SPCL*
Mean	80.11	632.57	637.39	4.12	4.06	4.28	1375.83	3390.45	1903.00
Med.	39.86	301.25	399.00	4.46	4.37	4.43	140.00	455.50	377.00
Max.	1791.89	6169.00	5999.95	5.00	5.00	5.00	35762.00	36093.00	35472.00
Min.	1.22	1.54	7.59	1.00	1.00	1.00	10	11.00	10
Std. Dev.	169.65	773.70	769.79	0.80	0.85	0.55	3271.14	6651.92	4685.68

Variable		*THICK*			*SFR*	
Category	*CONV*	*SHOP*	*SPCL*	*CONV*	*SHOP*	*SPCL*
Mean	13.77	12.88	9.37	0.14	0.17	0.14
Med.	10.00	10.00	6.00	0.13	0.13	0.00
Max.	42.00	33.00	29.00	1.00	1.00	1.00
Min.	2.00	2.00	2.00	0.00	0.00	0.00
Std. Dev.	10.66	8.70	7.97	0.16	0.20	0.20

3.4. Regression Model

To test the effects that the reputation and market variables have on sellers' prices, we set up a hierarchical regression model. The dummy variables *SF* and *TOP500* separate the effects on the specific seller types from the general effects. Moreover, only the markets with two or more active sellers are considered.

The moderated hierarchical regression model is presented in Equations (2.1) to (2.4). Due to standardization, the expected value of the dependent variable is equal to zero and therefore, the intercept is dropped from the regression equation. However, the unstandardized dummy intercepts are included to preserve their meaning. Equation (2.1) contains the basic additive model. We hypothesize that the ratings history moderates the rating score, which is captured by the interaction term $zRATE \cdot zHIST$. Deploying interactions into a regression equation require that nonlinear effects should be controlled for. As suggested by Cortina [1993], Equation (2.2) implements the quadratic effects as independent variables into the hierarchical model. The next step in Equation (2.3) includes two-way interactions in all variables presented in Equations (2.1) and (2.2). As a final step, the proposed moderator effects are added as two-way and three-way interactions in Equation (2.4).

$$zPRICE = \alpha_1 SF + \alpha_2 TOP + \beta_1 zRATE + \eta_1 zHIST + \lambda_1 zTHICK + \theta_1 zSFR \tag{2.1}$$

[15] We omit the fourth consumer goods category, unsought goods, because the data set does not include them.

$$+\varphi_1(zRATE)^2+\phi_1(zHIST)^2 \tag{2.2}$$

$$\begin{aligned}&+\beta_2 SF\cdot zRATE+\beta_3 TOP\cdot zRATE\\&+\varphi_2 SF(zRATE)^2+\varphi_3 TOP(zRATE)^2\\&+\eta_2 SF\cdot zHIST+\eta_3 TOP\cdot zHIST\\&+\phi_2 SF(zHIST)^2+\phi_3 TOP(zHIST)^2\\&+\lambda_2 SF\cdot zTHICK+\lambda_3 TOP\cdot zTHICK+\theta_2 SF\cdot zSFR+\theta_3 TOP\cdot zSFR\end{aligned} \tag{2.3}$$

$$\begin{aligned}&+\gamma_1 zRATE\cdot zHIST\\&+\gamma_2 SF\cdot zRATE\cdot zHIST+\gamma_3 TOP\cdot zRATE\cdot zHIST+\varepsilon\end{aligned} \tag{2.4}$$

The effect of a change in *zRATE* on the expected *zPRICE* is given by

$$\frac{\partial E[zPRICE\mid SF=0,TOP=0]}{\partial zRATE}=\beta_1+\varphi_1 zARATE+\gamma_1 zHIST \tag{3}$$

$$\frac{\partial E[zPRICE\mid SF=1,TOP=0]}{\partial zRATE}=\beta_1+\beta_2+(\varphi_1+\varphi_2)zARATE+(\gamma_1+\gamma_2)zHIST \tag{4}$$

$$\frac{\partial E[zPRICE\mid SF=0,TOP=1]}{\partial zRATE}=\beta_1+\beta_3+(\varphi_1+\varphi_3)zARATE+(\gamma_1+\gamma_3)zHIST\ . \tag{5}$$

If a nonlinear effect, which is assumed quadratic, exists, its magnitude depends on the level of the variable itself. The interpretation of interactions is very similar to this. Two-way interactions are usually a product of a dummy variable and a continuous variable giving a straightforward “on/off” interpretation. Interactions between two or more continuous variables mean that the effect of a change in the value of one covariate in an interaction term depends on the level of the other covariates that compose the interaction term.

As Equations (3) to (5) display, the effect decomposes to the sum of β_i coefficients (the linear main effect), the sum of φ_i coefficients multiplied by *zRATE* (the nonlinear main effect), and the sum of γ_i coefficients multiplied by *zHIST* (the interaction effect) where $i=1,2,3$ in all cases. If the nonlinear effect or the interaction effect is statistically significant, the effect of a change in *zRATE* on *zPRICE* is dependent on the level of *zRATE* and *zHIST*, respectively. In addition, the main effect should not be interpreted in isolation of the interaction effect[16]. Standardization of the variables makes the interpretation easier. The effect at the average length of the ratings history is equal to the main effect because the mean of a standardized covariate is zero.

4. Results

4.1. Heteroskedasticity, Multicollinearity and Moderation.

We estimate the hierarchical regression model using ordinary least squares regression (OLS). The results of OLS for Models 1 to 4 are reported in Table 5. The discussed numerical values refer to Model 4 unless indicated otherwise. Since each regression has problems with heteroscedasticity, we augment the estimates with the White’s heteroscedasticity consistent estimates. Furthermore, we run auxiliary regressions on each variable to detect multicollinearity (these regressions are omitted from this paper), which is often problematic in the regression models that utilize standardized variables. Variance inflation factors (VIF) for each model are obtained from auxiliary regressions. We report the range of VIFs in Table 5. They suggest that multicollinearity is not problematic with Models 1 and 2, but may cause some problems with Models 3 and 4. In Model 4, the VIFs range from 1.149 to 8.929. Since the magnitudes and statistical significance of estimated coefficients do not display wild variation as more variables are added, there is a good reason to believe that problematic multicollinearity remains muted.

In general, the addition of new variables increases the overall fit of the model because additional regressors naturally explain more of the variance of the dependent variable. Nevertheless, even Model 4 has a relatively low coefficient of determination (0.13), which means that the overall fit of the model is not very good. This is hardly surprising because many other factors affect a firm’s pricing decision. However, this does not preclude that estimates obtained from OLS are not good for analytical purposes, because the ceteris paribus relationship between the dependent variable and covariates does not depend directly on the magnitude of R^2 [Wooldridge 2002].

[16] Since it is difficult to interpret the interaction effects, no interpretations for $\frac{\partial E[zPRICE]}{\partial zHIST}$ are given.

The hierarchical regression procedure, which is described mathematically in Equations (2.1) to (2.4), proceeds as follows. First, Model 1 is a basic linear regression model. After this, Model 2 adds quadratic terms for the covariates that are hypothesized to interact with each other [Cortina 1993]. Then Model 3 introduces interactions by placing controls on the seller subgroups. Finally, Model 4 is the full model that includes all interactions. As the hierarchical model progresses by adding more covariates to the regression equation, the change in the coefficient of determination (ΔR^2) is small, but statistically significant (ΔF). This validates the inclusion of the additional regressors, and the hypothesis that a moderator effect is present.

From the moderation perspective, it is important to examine the type of moderation to determine whether the moderation effect addresses the strength or form of moderation. To find out the type of moderation, Sharma et al. [1981] propose a procedure for the identification of moderator variables. The procedure requires examining the statistical significance of the coefficients for *zHIST* which is the hypothesized moderator of *zRATE*. Since $\eta_1 = -0.080 \neq 0$ (the estimated coefficient for *zHIST*) in Model 1and $\gamma_1 = -0.117 \neq 0$ (the estimated coefficient for the interaction *zRATE*zHIST*) in Model 4, *zHIST* is a quasi-moderator. As a result, Sharma et al. [1981] suggest that the form of moderation is examined with different values of the moderator variable. If *zHIST* influences the form of moderation, the use of the moderated regression analysis is warranted [Venkatraman 1989].

Examining the results obtained from Model 4 gives insights of how a seller's reputation and the market environment influence prices. Estimated coefficients indicate how much a change of one standard deviation in a covariate alters the dependent variable. Since interactions are present in the regression equation, estimates for the standalone term *zRATE* must be assessed when the moderator, *zHIST*, equals zero. Conveniently, this coincides with the mean as explained earlier. Standardization provides an interpretation for the deployed dummy variables. Only the dummy for storefronts (0.502) is statistically significant, whereas the dummy for Top500-sellers is not. Therefore, when all variables are set to their mean values, the storefront prices are a half standard deviation higher than the prices of other sellers, but Top500-sellers do not differ from other sellers. In consequence, the dummy variables fail to support both H2A and H2B.

4.2. Reputation Variables

The signs of the estimated coefficients of *zRATE* are statistically significant and negative for all sellers (-0.111) and storefronts (-0.335). This implies an inverse relation between prices and rating scores when *zRATE* and *zHIST* are set equal to zero. The quadratic effect of $zRATE^2$ is negative (-0.051) for all sellers, but positive forTop500-sellers (0.099). As a result, the main effect of the rating score is curvilinear. For the general seller population and storefronts, these results indicate a strong inverse relation between prices and rating scores. However, the estimated coefficient of $TOP500*zRATE^2$ for is positive and greater by magnitude than $zRATE^2$. As a consequence, high rating scores dominate the general first-order and second-order effects providing returns on reputation for Top500-sellers.

Table 5: OLS Estimates for Hierarchical Regression Model

Variable	***Model 1***	***Model 2***	***Model 3***	***Model 4***
SF	**0.543*** (0.031)**	**0.565*** (0.032)**	**0.507*** (0.052)**	**0.502*** (0.052)**
TOP500	**0.084*** (0.020)**	**0.127*** (0.022)**	0.026 (0.027)	0.029 (0.027)
zRATE	**-0.147*** (0.012)**	**-0.171*** (0.013)**	**-0.080*** (0.016)**	**-0.111*** (0.016)**
zHIST	**-0.080*** (0.010)**	**-0.058*** (0.016)**	**-0.089*** (0.019)**	**-0.061*** (0.021)**
zTHICK	**-0.027*** (0.009)**	**-0.018* (0.009)**	**-0.079*** (0.012)**	**-0.076*** (0.012)**
zSFR	**-0.127*** (0.011)**	**-0.126*** (0.011)**	**-0.051*** (0.016)**	**-0.046*** (0.016)**
$(zRATE)^2$		**-0.028*** (0.009)**	**-0.068*** (0.015)**	**-0.051*** (0.015)**
$(zHIST)^2$		**-0.015** (0.007)**	-0.012 (0.008)	-0.000 (0.008)
SF*zRATE			**-0.446*** (0.040)**	**-0.335*** (0.046)**
SF*zHIST			-0.016 (0.047)	-0.064 (0.048)
SF*zTHICK			**0.188*** (0.030)**	**0.183*** (0.030)**
SF*zSFR			**-0.094*** (0.028)**	**-0.092*** (0.028)**

SF*(zRATE)²			-0.019 (0.036)	-0.051 (0.037)
SF*(zHIST)²			0.002 (0.036)	-0.043 (0.036)
TOP500*zRATE			0.024 (0.036)	0.056 (0.036)
TOP500* zHIST			-0.022 (0.043)	-0.049 (0.043)
TOP500*zTHICK			**0.036* (0.021)**	0.031 (0.021)
TOP500*zSFR			**-0.077** (0.030)**	**-0.082*** (0.030)**
TOP500*(zRATE)²			**0.106*** (0.022)**	**0.099*** (0.024)**
TOP500*(zHIST)²			**0.046** (0.019)**	**0.047** (0.021)**
zRATE*zHIST				**-0.117*** (0.025)**
SF*zRATE*zHIST				**0.341*** (0.065)**
TOP500*zRATE*zHIST				0.065 (0.048)
Observations	8916	8916	8916	8916
R^2	0.092	0.094	0.127	0.130
F	181.524***	132.621***	68.041***	60.646***
ΔR^2		0.002	0.033	0.003
ΔF		7.788***	21.037***	9.211***
VIF Range	0.957 – 1.233	1.104 – 2.421	1.145 – 7.407	1.149 – 8.929
*** p-value<0.01; ** p-value<0.05; * p-value<0.1; Standard errors in parentheses.				

The estimated coefficient of *zHIST* is statistically significant and negative (-0.061), but the coefficients for the control groups are statistically insignificant. There is no quadratic effect on the general seller population which suggests that the overall influence of the ratings history is linear. Thus, longer ratings histories are associated with lower relative prices. The quadratic effect is statistically insignificant also for storefronts, but the estimated coefficient of $TOP500*zHIST^2$ is positive (0.047) and statistically significant. Hence, the effect of the ratings history on Top500-sellers is nonlinear implying that higher prices have a positive correlation with the longer ratings histories.

The interaction terms are statistically significant for all sellers and storefronts. The overall effect (-0.117) is negative, whereas the estimate for storefronts (0.341) is positive. As a consequence, the combined effect of the interaction is positive on storefronts indicating returns on reputation for longer ratings histories. In contrast, longer ratings histories are associated with lower prices among other sellers.

These results are further explored by testing Model 4 with different values of the ratings history and rating score. These values are entered into Equations (3) to (5) and assuming that the market variables are equal to zero. Tables 5 to 7 lay out the total effects on the price while Figures 1 to 3 illustrate these results. The values used for *zHIST* and *zRATE* are the minimum (min.), mean, median (med.) and maximum (max.) which are obtained from Table 3.

Table 6 together with Figure 1 demonstrates the total effect on all sellers. The graphical representation shows that there is an overall trend that better reputations correlate with lower prices. In addition, longer ratings histories correlate with lower prices. The difference between the two extremes is more than 100% (a full standard deviation), as the minimum rating score and ratings history is 0.286, while the maximum values produce -0.736. In consequence, there seems to be no returns on reputation when all sellers are considered, because all positive total effects can be found in the minimum *zRATE* column and with the ratings histories at the mean or below. Thus, the evidence does not support either H1A or H1B when all sellers are considered.

Both controlled seller groups display results that differ from the general seller population. Table 7 and Figure 2 present the effects on storefronts. There is again a visible negative trend between the rating score and prices. In contrast, the longer ratings histories are related to the higher prices. The only positive coefficients can be found in the row of the maximum length of *zHIST*. Thus, an increase in *zRATE* by one standard deviation increases the price by 0.788 to 0.466 standard deviations when *zHIST* is at the maximum. As with the whole sample, the difference between the minimum values (-0.581) and the maximum values (0.466) is again more than 100%. While rating scores are related to lower prices, these figures show that the longer ratings histories may give some pricing power to storefronts. In conclusion, we find support for H1B but no support for H2B when storefronts are considered.

Table 6: Total Effect on All Sellers with Different Values of *zRATE* and *zHIST*

zHIST	zPRICE			
max.	-0.424	-0.639	-0.653	-0.736
med.	0.150	-0.065	-0.079	-0.162
mean	0.104	-0.111	-0.125	-0.208
min.	0.286	0.072	0.057	-0.026
zRATE	min.	mean	med.	max.

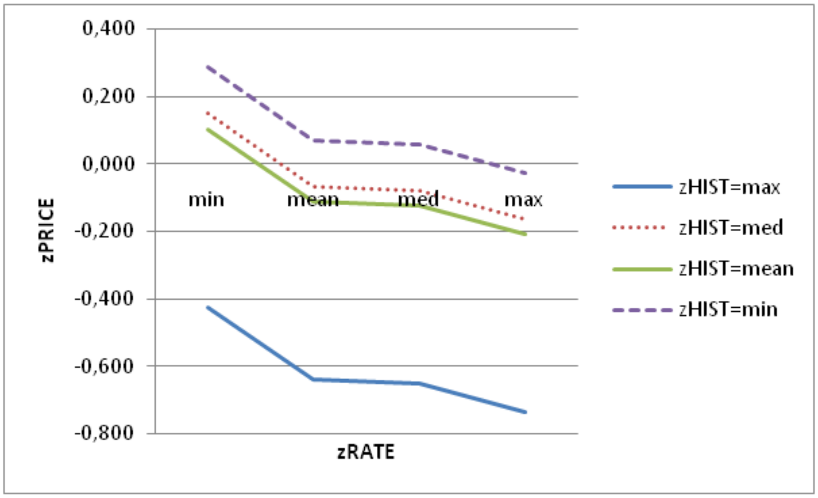

Figure 1: Total Effect on All Sellers with Different Values of *zRATE* and *zHIST*

Table 7: Total Effect on Storefronts with Different Values of *zRATE* and *zHIST*

zHIST	SF*zPRICE			
max.	0.778	0.564	0.549	0.466
med.	-0.320	-0.534	-0.549	-0.632
mean	-0.232	-0.446	-0.461	-0.544
min.	-0.581	-0.796	-0.810	-0.893
zRATE	min.	mean	med.	max.

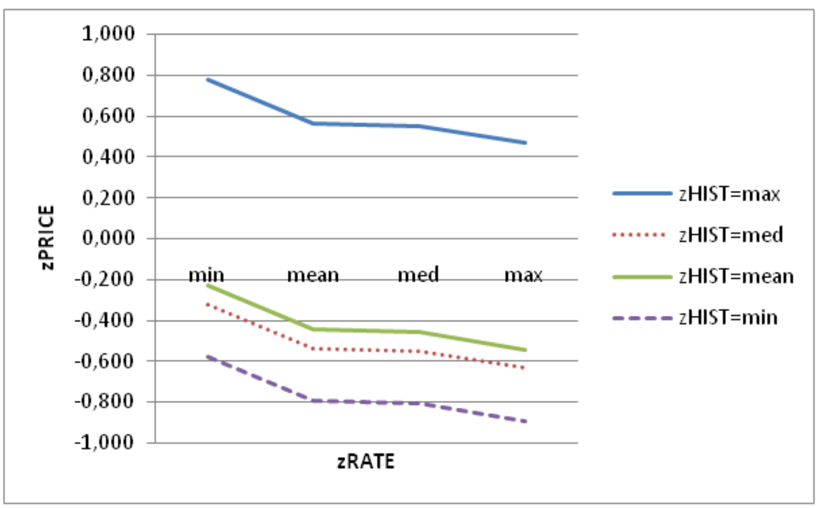

Figure 2: Total Effect on Storefronts with Different Values of *zRATE* and *zHIST*

The results for Top500-sellers are shown in Table 8 and Figure 3. Unlike other sellers, there is a visible positive relation between prices and rating scores. Although most values are in the negative territory, their magnitude increases towards the higher rating scores. Nevertheless, the longer ratings histories correlate with the lower prices. The difference between the minimum values (-0.129) and the maximum values (-0.547) is considerably lower at about 40%. Positive values can be found in the minimum row of *zHIST* at and above the mean *zRATE*, and at the median *zHIST* and the maximum *zRATE* cell. The magnitudes of these effects, however, are not sizable corresponding to 0.027 to 0.164 standard deviations. Nonetheless, these figures provide some support for H1A and H2A when Top500-sellers are considered.

Table 8: Total Effect on Top500-Sellers with Different Values of *zRATE* and *zHIST*

zHIST	TOP500*zPRI CE			
max.	-0.840	-0.639	-0.625	-0.547
med.	-0.266	-0.065	-0.051	0.027
mean	-0.312	-0.111	-0.097	-0.019
min.	-0.129	0.072	0.086	0.164
zRATE	min.	mean	med.	max.

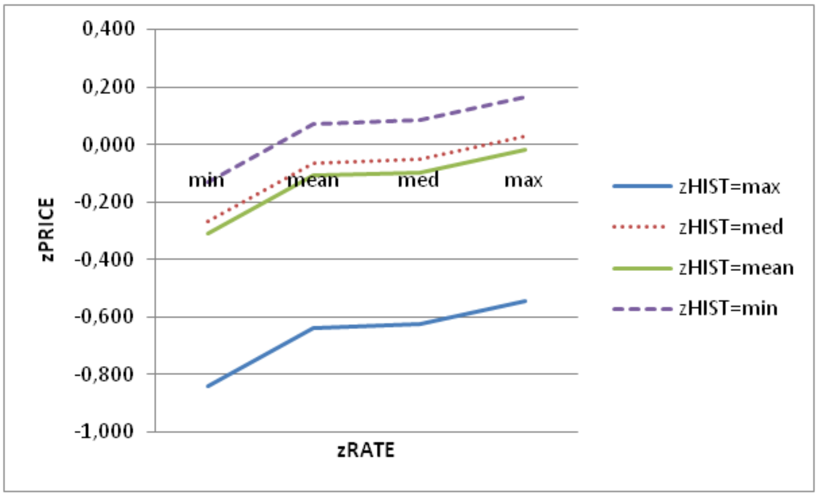

Figure 3: Total Effect on Top500-Sellers with Different Values of *zRATE* and *zHIST*

4.3. Market Variables

The coefficients for market variables, *zTHICK* and *zSFR*, provide straightforward interpretations. The coefficient for *zTHICK* (-0.076) is negative and statistically significant while the estimate for Top500-sellers is not. However, the estimated coefficient for storefronts (0.183) is positive and greater in magnitude. Thus, sellers in general decrease prices when the number of market incumbents increases. Storefronts, on the other hand, realize price gains as the increased competition drives prices towards a competitive price. Based on this evidence, H3A cannot be rejected entirely. The storefront ratio has a stronger impact on prices. Estimates are negative for all sellers (-0.046), storefronts (-0.092) and Top500-sellers (-0.082). This indicates that an increase in the proportional number of storefronts in a market has an adverse effect on prices. In addition, this effect is more pronounced among the control groups. As a result, H3B is accepted.

4.4. Quantile Regression

Quantile Regression (QR) is a semiparametric estimation method. The strength of QR is that it is robust to outliers in data [Koenker and Hallock 2001]. A special case of QR is median regression which is a semiparametric alternative to OLS. As the name suggests, median regression reveals only a fraction of the information that QR can give. With QR, it is possible to get estimates for coefficients across the conditional distribution of the dependent variable. We use QR to check how robust the results of OLS in Model 4 are across the conditional distribution. The quantile regression estimates are reported in Table 9. They are provided for the first quartile (0.25 quantile), the second quartile i.e. the median (0.50 quantile) and the third quartile (0.75 quantile). The first quartile corresponds to the lower tail of the dependent variable's distribution and the third quartile corresponds to the upper tail.

Table 9: Quantile Regression Estimates for Model 4

Variable	***0.25***	***0.5***	***0.75***
SF	**-0.437*** (0.070)**	**0.472*** (0.085)**	**1.391*** (0.068)**
TOP500	**-0.555*** (0.034)**	**0.077** (0.039)**	**0.054*** (0.034)**
zRATE	**-0.335*** (0.027)**	**-0.175*** (0.025)**	0.115 (0.029)
zHIST	**0.222*** (0.035)**	-0.054 (0.041)	**-0.280*** (0.034)**
zTHICK	**-0.135*** (0.011)**	**-0.132*** (0.013)**	**-0.134*** (0.014)**
zSFR	**0.034* (0.020)**	**-0.077*** (0.025)**	**-0.125*** (0.025)**
(zRATE)²	**-0.450*** (0.038)**	**-0.130*** (0.021)**	**0.278*** (0.035)**
(zHIST)²	**-0.237*** (0.032)**	0.017 (0.012)	**0.013*** (0.017)**
SF*zRATE	**-0.107* (0.064)**	**-0.418*** (0.069)**	**-0.453*** (0.066)**
SF*zHIST	**-0.392*** (0.062)**	-0.075 (0.072)	**0.305*** (0.059)**
SF*zTHICK	**0.262*** (0.044)**	**0.195*** (0.050)**	**0.199*** (0.034)**
SF*zSFR	**-0.069** (0.034)**	**-0.076* (0.040)**	**-0.094** (0.038)**
SF*(zRATE)²	**0.377*** (0.062)**	0.062 (0.052)	**-0.037*** (0.051)**
SF*(zHIST)²	**0.187** (0.075)**	-0.006 (0.062)	**-0.268*** (0.070)**
TOP500*zRATE	**0.292*** (0.046)**	0.071 (0.056)	-0.052 (0.050)
TOP500* zHIST	**-0.248*** (0.053)**	-0.022 (0.075)	**0.135** (0.059)**
TOP500*zTHICK	**0.167*** (0.020)**	**0.0486* (0.028)**	-0.000 (0.028)
TOP500*zSFR	**-0.114*** (0.030)**	**-0.163** (0.043)**	-0.021 (0.048)
TOP500*(zRATE)²	**0.486*** (0.046)**	**0.141*** (0.033)**	**-0.197*** (0.042)**
TOP500*(zHIST)²	**0.245*** (0.036)**	**-0.000 (0.032)**	**-0.067** (0.027)**
zRATE*zHIST	-0.060 (0.055)	**-0.133*** (0.045)**	**-0.155*** (0.050)**
SF*zRATE*zHIST	**0.342*** (0.088)**	**0.276*** (0.099)**	**0.394*** (0.086)**
TOP500*zRATE*zHIST	-0.037 (0.068)	0.112 (0.076)	**0.136* (0.076)**
Observations	8916	8916	8916
Pseudo R^2		0.070	
*** p-value<0.01; ** p-value<0.05; * p-value<0.1; Standard errors in parentheses.			

The dummy variables behave as expected showing negative values in the lower tail and positive values in the upper tail. The estimates verify that when all values are zero at the mean, storefronts set higher prices than other sellers. For example, compared with Top500-sellers (0.541), storefront prices (1.391) are considerably higher in the upper tail. The estimates for coefficients of the market variables as described in Equations (3) to (5) are combined to Equations (6) to (8), so the results are easier to interpret. The values in parentheses are estimates for each quartile, the first quartile on the top, the median in the middle, and the third quartile at the bottom.

Setting the levels of *zRATE* and *zHIST* equal to zero indicates that the negative effect of reputation dominates almost everywhere. Nevertheless, varying *zRATE* sufficiently, as *zHIST* is held constant, brings positive effects for Top500-sellers at and above the median. However, z*RATE* must be very large to offset the negative estimates of β at the median. Contrary to OLS estimates, all sellers may enjoy returns on reputation in the upper tail for all values of *zRATE* that exceed the mean. In contrast, the negative relationship between the *zRATE* and *zPRICE* persists with storefronts. Also, we find negative effects on all sellers at and below the median and in the lower tail for Top500-sellers. As a result, QR estimates show that the predictions of OLS hold everywhere only for storefronts but not for Top500-sellers and the general seller population.

$$\frac{\partial E\left[zPRICE \mid SF=0, TOP500=1\right]}{\partial zRATE}=\begin{pmatrix}-0.335\\-0.175\\0\end{pmatrix}+\begin{pmatrix}-0.450\\-0.130\\0.278\end{pmatrix}zRATE+\begin{pmatrix}0\\-0.133\\-0.155\end{pmatrix}zHIST \tag{6}$$

$$\frac{\partial E\left[zPRICE \mid SF=1, TOP500=0\right]}{\partial zRATE}=\begin{pmatrix}-0.432\\-0.593\\-0.413\end{pmatrix}+\begin{pmatrix}-0.073\\-0.130\\-0.088\end{pmatrix}zRATE+\begin{pmatrix}0.342\\-0.143\\0.239\end{pmatrix}zHIST \tag{7}$$

$$\frac{\partial E\left[zPRICE \mid SF=0, TOP500=1\right]}{\partial zRATE}=\begin{pmatrix}-0.043\\-0.175\\-0.052\end{pmatrix}+\begin{pmatrix}-0.036\\0.011\\0.081\end{pmatrix}zRATE+\begin{pmatrix}0\\-0.133\\-0.019\end{pmatrix}zHIST \tag{8}$$

Holding *zRATE* constant and varying *zHIST* shows that the predictions of OLS are not equal all over the distribution. While the estimates for the median and upper tail are negative for the general seller population and Top500-sellers, estimates for the lower tail are statistically insignificant. This implies that the negative effect of an increase in *zHIST* does not affect the sellers who charge the lowest prices. In addition, OLS predicted that storefronts, on average, respond positively to an increase in the length of ratings history. QR estimates show that the effect is negative at the median but positive in both tails. Therefore, we can conclude that an increase in *zHIST* benefits the sellers whose prices are in the lower tail and upper tail of the price distribution.

Estimates for the market variables, *zTHICK* and *zSFR*, display mixed results. The estimated coefficients for *zTHICK* are negative for all sellers at all quartiles (-0.135; -0.132; -0.138). This effect is strong enough to offset the positive coefficients for Top500-sellers at the median (0.048) but not in the lower tail (0.167) while the upper tail is not significant. Storefronts, however, show a positive response everywhere (0.262; 0.195; 0.199) which offsets the negative influence of the overall estimates. Hence, increasing market thickness leads to lower prices but the sellers in the lower tail realize relative gains as the prices fall more in the upper tail of the price distribution. On the other hand, the influence of the storefront ratio is negative for storefronts everywhere (-0.069; -0.076; -0.094) which offset and accentuate the overall influence of *zSFR* (0.034; -0.077; -0.125). In a similar manner, the prices of Top500-sellers respond negatively to an increase in the storefront ratio at and below the median (-0.114; -0.163). Overall, these results imply that intensified competition depresses prices more in the upper tail of the price distribution.

4.5. Product Type

The results of regressions for different product types are presented in Table 10[17]. The lack of statistical significance in estimated coefficients is striking. An explanation for this could be multicollinearity which inflates variances. The auxiliary regressions indicate that multicollinearity may be problematic especially with convenience goods and shopping goods because some VIFs exceed the generally accepted tolerance level of 10.

In the convenience goods category, only the coefficient for *zRATE*zHIST* is statistically significant (-0.342). This indicates that there are no returns on reputation in general. For storefronts, the main effect (-0.407) and the interaction (0.777) are statistically significant, and the combined coefficients for the interaction terms sum to 0.435. So only ratings histories that are significantly above the mean provide returns on reputation because they have to offset the negative main effect. The combined coefficients of the interaction terms sum to 0.362 for Top500-sellers, so any ratings histories above the mean could enable price premiums. As regards the market variables, only the storefront ratio is statistically significant and negative (-0.121). This implies that an increase in the number of storefronts intensifies competition in the markets for convenience goods.

In the shopping goods category, only the interaction term is statistically significant (-0.427) for all sellers. Thus, there are no returns on reputation for the general seller population. The combined coefficients for storefronts sum to 0.279. Thus histories significantly above the mean could enable premium pricing. For Top500-sellers, however, the

[17] Only estimates for Model 4 are reported. Estimates for Models 1 to 3 are available from the author upon request.

sum is slightly below zero (-0.073), which suggests that no returns on reputation exist. The market variables display diverse effects. An increase in market thickness (-0.101) tends to lower prices, but this does not affect storefronts, for which the combined effect is positive (0.118). So their relative price position improves as market thickness increases. The storefront ratio is negative and statistically significant for storefronts (-0.274) and Top500-sellers (-0.349). This suggests that storefronts intensify competition in the shopping goods markets.

Table 10: OLS Estimates for Different Product Categories

Variable	***CONV***	***SHOP***	***SPCL***
SF	**0.567 (0.157***)**	**0.921 (0.219***)**	**0.787 (0.244***)**
TOP500	-0.081 (0.102)	0.062 (0.070)	0.0.3 (0.094)
zRATE	-0.061 (0.065)	-0.040 (0.059)	**-0.125 (0.060**)**
zHIST	-0.040 (0.077)	-0.080 (0.079)	-0.008 (0.077)
zTHICK	-0.036 (0.039)	**-0.101 (0.035***)**	**-0.128 (0.040***)**
zSFR	**-0.121 (0.064*)**	0.053 (0.053)	-0.066 (0.060)
(zRATE)^2	0.013 (0.067)	-0.003 (0.030)	-0.069 (0.053)
(zHIST)^2	0.010 (0.027)	0.030 (0.300)	-0.009 (0.032)
SF*zRATE	**-0.407 (0.174**)**	**-0.647 (0.168***)**	-0.239 (0.232)
SF*zHIST	**0.369 (0.160**)**	0.261 (0.189)	0.109 (0.158)
SF*zTHICK	-0.034 (0.079)	**0.219 (0.093**)**	**0.607 (0.178***)**
SF*zSFR	-0.009 (0.095)	**-0.247 (0.088***)**	-0.043 (0.131)
SF*(zRATE)^2	-0.139 (0.117)	-0.125 (0.120)	0.019 (0.181)
SF*(zHIST)^2	-0.092 (0.178)	-0.393 (0.261)	**-0.328 (0.136**)**
TOP500*zRATE	-0.073 (0.155)	-0.049 (0.102)	-0.173 (0.119)
TOP500* zHIST	-0.199 (0.193)	0.086 (0.134)	-0.159 (0.125)
TOP500*zTHICK	0.114 (0.070)	-0.017 (0.066)	-0.010 (0.008)
TOP500*zSFR	0.000 (0.124)	**-0.349 (0.089***)**	**-0.204 (0.108*)**
TOP500*(zRATE)^2	-0.029 (0.096)	-0.009 (0.061)	0.034 (0.079)
TOP500*(zHIST)^2	-0.033 (0.079)	-0.003 (0.057)	0.073 (0.076)
zRATE*zHIST	**-0.342 (0.108***)**	**-0.427 (0.107***)**	-0.084 (0.090)
SF*zRATE*zHIST	**0.777 (0.274***)**	**0.706 (0.252***)**	0.286 (0.249)
TOP500*zRATE*zHIST	**0.704 (0.200***)**	**0.354 (0.156**)**	0.080 (0.156)
Observations	838	1055	581
R^2	0.170	0.238	0.269
F	7.590***	14.679***	9.314***
VIF Range	1.754-17.857	1.783-15.625	1.742-6.803
*** p-value<0.01; ** p-value<0.05; * p-value<0.1; Standard errors in parentheses.			

In the specialty goods category, all statistically significant estimates for reputation variables are negative. Hence, there are no returns on reputation in this product category. The results are inconclusive for the market variables. Overall, an increase in market thickness lowers prices, but the combined effect is positive for storefronts (0.479). However, these results are most likely caused by the low frequency of storefronts in the specialty goods markets. The storefront ratio is statistically significant for Top500-sellers (-0.204), which could mean that small firms may force well-known sellers into price competition.

In conclusion, the influence of the product type is rather unexpected. Returns on reputation, when available, are the greatest in the markets for convenience goods. In contrast, they do not exist in the markets for specialty goods while the markets for shopping goods show some possibilities for price premiums. For this reason, H4 is rejected.

5. Discussion

5.1. Key Findings

The objective of this study was to determine whether the returns on reputation as price premiums exist in online retail markets. The extant research on online auctions has verified that sellers with better reputations as measured by their reputation scores earn returns on reputation. In this study, we analyze data that is obtained from the reputation system of the online price comparison website Pricegrabber.com.

We use a hierarchical regression model to answer the research hypotheses laid out in Section 2.4. When all sellers are considered, the results do not support the hypothesis that sellers earn returns on reputation. On the contrary, the evidence points to the opposite effect. Higher rating scores and longer ratings histories correlate with lower prices.

We hypothesized also that branded firms (Top500-sellers) may earn returns to their brands, which are in part manifestations of reputation, whereas small enterprises (storefronts) compete with price. The results indicate that without reputation effects, this is not the case. The prices set by Top500-sellers do not differ from the general seller population, whereas storefronts set higher prices than other sellers. However, the estimated total effects indicate that Top500-sellers may earn some returns on their reputation, but storefronts do not. In addition, storefronts may receive a small competitive advantage from sufficiently long ratings histories. Quantile regression indicates that the general seller population may earn returns on reputation in the upper tail of the price distribution. In contrast, there is a positive correlation between higher prices and longer ratings histories in the storefront group, but Top500-sellers display the opposite effect. In conclusion, Top500-sellers and storefronts provide some support for the hypothesis on returns on reputation as measured by reputation scores and the length of ratings history, respectively.

As a robustness check, we test whether the product type has any influence on price premiums. Three broad product categories (convenience goods, shopping goods and specialty goods) are being analyzed. The theory suggests that consumers perceive specific and valuable goods to be riskier purchases than convenience goods This could give an opportunity to benefit from a good reputation. However, our findings indicate that this is not the case in retail e-commerce. Storefronts and Top500-sellers may earn returns on reputation in convenience goods markets which are the lowest risk markets for consumers. In contrast, returns on reputation are vague in shopping goods markets and they do not exist in specialty goods markets.

The theory of industrial organization suggests that an increase in the number of market incumbents intensifies price competition. We hypothesize that increases in market thickness and in the relative number of storefronts in a market (the storefront ratio) leads to lower prices. Again, the findings are inconclusive. Our study finds that an overall increase in market thickness lowers prices. Quantile regression verifies that the effect is the strongest in the upper tail of the price distribution. However, Top500-sellers show no response to an increase in market thickness while storefront prices increase relative to other prices with market thickness. A more influential variable in intensifying price competition seems to be the storefront ratio.

5.2. Suggested Explanations for Findings

Compared with the online auctions literature, we fail to detect an overall positive relationship between seller reputations and price premiums. An explanation for this could be that rating scores are a crude measure of a seller's reputation. Thus, unlike in auction markets, buyers can find additional information about sellers elsewhere and base their purchase decisions on this information. The relationship between lower prices and longer ratings histories suggests that ratings histories may proxy the level of sales also in retail markets as suggested by Cabral and Hortaçsu [2006]. As a consequence, comparison shopping services fulfill their purpose as a means to conduct price comparisons. Low prices attract more buyers who are satisfied with the low prices and return positive feedback. As control groups diverge from the general population, the seller type could explain results. First, Top500-sellers are well-known sellers. As a result, the sellers with better reputations within the group receive some returns on reputation. However, the negative correlation between the length of ratings history and prices imply that low prices attract more buyers. Second, storefronts may attempt to gain occasional sales at high prices which would explain the dummy variable that indicates a price level that is higher than the average price level of other sellers. The positive effect of the ratings history could mean that it signals a commitment to stay in the market. This could be important to buyers because returns and warranties often require contacting the seller.

Another explanation for the differences could be that the value of merchandise determines the importance of a seller's reputation in a purchase decision. Consumers may view a reputation as an insurance against fraudulent behavior when purchasing valuable items such as televisions. As the value of a purchased item increases, consumers become more risk averse and are willing to pay a price premium for a homogenous item to a more reputable seller.

The price data provides some support for this proposition. The descriptive statistics of prices are presented in Table 11. They show that the mean (279.90) and range (8449.01) of storefronts prices is considerably less than the mean and range for the other seller types. The difference in the median prices is less severe. The statistical tests in Table 12 support the argument that on average, storefronts focus on items that are of lower value, because the mean and median prices are lower than those of the other two seller types. While the means of Top500-sellers and other sellers are equal, their medians are not. This implies that Top500-sellers sell more valuable items than other sellers, but this difference is not very large.

This explanation is not entirely convincing because a sub-sample of product types reveals an inverse trend. Returns on reputation are the most apparent among convenience goods which are the least expensive on average (see Table 4). In contrast, price premiums wane in the shopping goods markets and disappear in the specialty goods markets. Furthermore, measured by market thickness, the markets for convenience goods are more competitive either the markets for shopping goods or specialty goods. An explanation for these findings could be consumer search which is more intense for shopping goods and specialty goods [Thirumalai and Sinha 2009]. Buyers use comparison shopping websites to locate the best deals for the desired goods. In consequence, this reduces a seller's ability to benefit from a good reputation.

The disparate impact of market thickness is more difficult to explain. As market thickness increases, the likelihood that a consumer encounters a seller that has the best fit for her preferences increases because the number of firms in the market increases. For example, earlier successful transactions with a familiar seller, or a seller's brand could weigh in consumer decision making which could explain the effects on Top500-sellers and storefronts. According to Grover et al. [2006], information overload – too much information to process – might cause consumers to buy only at known firms.

Table 11: Descriptive Statistics of Prices of Seller Groups

Statistic	***Other Sellers***	***Storefronts***	***Top500-Sellers***
Mean	423.89	279.90	426.07
Median	111.12	85.00	149.99
Maximum	16797.59	8450.00	16265.20
Minimum	0.01	0.99	0.93
Std. Dev.	962.81	535.48	881.98
Observations	9681	2173	2353

Table 12: T-Tests and Mann-Whitney Tests for Seller Groups

Hypothesis	***T-test***	***Mann-Whitney***	***Conclusion***
Storefronts =Top500-sellers	-6.673***	9.512***	Rejected / Rejected
Storefronts =Other Sellers	-6.741***	5.054***	Rejected / Rejected
Top500-sellers =Other Sellers	-0.100	6.145***	Accepted / Rejected

*** p-value<0.01; ** p-value<0.05; * p-value<0.1.

5.3. Conclusion

In this paper, we examined the effect of a seller's reputation on pricing in retail e-commerce. In our model, a seller's reputation is measured by its rating score and ratings history. This study used data from Pricegrabber.com which is a popular comparison shopping service. We control for two seller types, storefronts and Top500-sellers, based on their level of sales and choice over a sales channel in Pricegrabber.com. We also distinguish between convenience, shopping and specialty goods. Moreover, the variables that characterize competition in markets (market thickness and the storefront ratio) were included in the regression model. OLS and quantile regression with standardized variables is used to estimate a hierarchical regression model.

Our findings indicate that when all sellers are considered, there are no universal returns on reputation in retail e-commerce. However, the specified seller groups display some returns on the measures of reputation. Top500-sellers show returns on reputation as measured by the rating score while the longer ratings histories allow premium pricing for storefronts. Quantile regression indicates that returns on reputation concentrate to the upper quartile of the price distribution, and even the general seller population may obtain price premiums. The impact of competition on prices proves mixed. Overall, an increase in market thickness lowers prices as expected. In contrast, the results indicate that Top500-sellers are not affected by this increase, and storefront prices increase with market thickness. However, an increase in the storefront ratio lowers prices in all seller groups.

5.4. Contribution

The main contribution of this paper is to expand the empirical research of reputation effects in e-commerce from online auctions to online retail markets. The influence of reputation scores on prices has not been studied in this scale in the context of retail e-markets. Therefore, this study fills the void that has existed in e-commerce research and contributes to the extant literature of reputation effects in e-commerce.

Overall, these results show that a reputation score does not enable price premiums when all sellers are considered. Unfortunately, our results are not directly comparable to those obtained from auction markets because of the fundamental differences between retail and auction markets. The closest comparison is Baylis and Perloff [2002] who find that favorable third party ratings have no effect on prices. They conclude that there is no premium associated with the ratings because “bad firms” charge higher prices than “good firms”. Our study points to the same direction. However, we hesitate to give this interpretation because the average rating score is over 4 in a scale of 1 to 5. Consequently, “bad firms” are not abundant in the market. In addition, it is impossible to tell how many sales these firms are able to conclude after bad reviews start to accumulate. In the end, even small differences in rating scores and ratings histories could matter if sellers were otherwise similar.

Our evidence shows that Top500-sellers earn returns on their investments in reputation. One explanation could be brand recognition as suggested by Brynjolfsson and Smith [2001]. On the other hand, by definition Top500-sellers are major players in e-commerce. Running an e-commerce operation with such magnitude requires investments in state-of-the-art technology and good services [Saloner & Borenstein 2001]. Therefore, it is not surprising that returns on reputation are found in this seller group. As a result, investment in a large-scale, well-run e-commerce operation may result in a good reputation that, in turn, is a competitive advantage in retail e-commerce.

Storefronts displayed returns on the length of ratings history. As these sellers are small-scale e-commerce vendors, this suggests that a consistent ratings history signals commitment to stay in the market to buyers. Buyers could value experience which is reported by McDonald and Slawson [2002] in online auction markets. They find that experienced sellers receive more bids and end up receiving higher final prices. These findings suggest that a small-scale e-commerce start-up should sell at low prices initially to generate and build up a consistent ratings history. Eventually, this strategy may allow charging a price premium over other sellers.

Our evidence indicates also that the number of sellers in the market have an impact on pricing. More precisely, the number of storefronts seems to affect more than the overall number of sellers. The seller type could explain this. As storefronts sell their products through the comparison shopping website, they have a larger propensity to enter into price competition. For other sellers, a comparison shopping website is a way to attract more price-conscious customers because they may derive a large bulk of sales elsewhere. For this reason, they have no interest in entering price competition for the informed consumers who look for bargains. Instead, the market for them resembles the situations described in Salop and Stiglitz [1977] and Varian [1980] where the informed customers pay lower prices and the uninformed pay higher prices. Occasionally, a firm charges the lowest price universally by organizing a sale or just by being lucky.

5.5. Implications

The findings of this study provide information about the importance of a seller’s reputation in highly competitive online comparison shopping markets. For small enterprises, a commitment to stay in the market, which is reflected by the length of a firm’s ratings history, may provide some pricing power. Large, e-commerce enterprises seem to reap some benefits from their reputations. In light of this evidence, a firm should not overestimate the competitive advantage of a good reputation. It is likely that as long as a seller does not deviate from its peers significantly, buyers overlook small differences in reputations and prefer lower prices. Thus, active reputation building may not be an effective tool in building a competitive advantage in comparison shopping markets, whose main purpose is to enable consumers’ price search across different vendors. In contrast, the low price strategy, at least initially, may be effective for small enterprises to gain consumer trust. As small firms become established sellers, they may be able to set small price premiums over market entrants.

5.6. Limitations and Suggestions for Future Research

There are a few obvious problems with this study. First, the overall fit of the model is not very good. To address this problem, the scope of the study should probably be narrowed down to include a smaller number of products. Second, the sample of sellers may be biased towards sellers whose strategy is to compete with price. This is because buyers who use a comparison shopping service are likely to be more experienced users of e-commerce and more price sensitive than the average buyer. As a result, sellers who choose to participate in a comparison shopping market choose to compete with price and not with service quality, for instance. Third, reputation systems provide more information than numerical scores in the form of written comments. Analyzing these comments, and whether or not they have an impact on transactions, opens a new avenue to research as demonstrated by Pavlou and Dimoka [2006]. Fourth, this research, as well as many other empirical inquiries into e-commerce, use posted prices. For this reason, it is impossible to tell which offers lead to concluded transactions. Access to the data that shows details about concluded transactions would greatly enhance knowledge about the influence of a seller's characteristics in e-commerce. Indeed, data from e-commerce companies could provide insight how much a seller's efforts to increase buyer's switching costs through loyalty programs mitigates consumer's price sensitivity and what is the influence of the seller's reputation in that case.

Acknowledgement

The author wishes to thank the Yrjö Jahnsson Foundation, the North Karelia Regional Fund and the Faculty of Law, Economics and Business Administration of University of Joensuu for the financial support and the anonymous reviewers for insightful comments.

REFERENCES

Akerlof, George A., "The Market for Lemons: Quality Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, 84 (3):488-500, 1970.

Ba, Sulin & Pavlou, Paul A., "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior", *MIS Quarterly*, Vol. 26, No. 3:243-268, 2002.

Bakos, Yannis & Dellarocas, Chrysanthos, "Cooperation without Enforcement? A Comparative Analysis of Litigation and Online Reputation as Quality Assurance Mechanism", *MIT Sloan Working Paper*, No. 4295-03, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=393041, 2003.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior", *The Quarterly Journal of Economics*, Vol. 107, No. 3:797-817, 1992.

Baye, Michael R., Morgan, John & Scholten, Patrick, "Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site", *The Journal of Industrial Economics*, Vol. 52, No. 4:463-496, 2004.

Baylis, Kathy & Perloff, Jeffrey M., "Price Dispersion on the Internet: Good Firms and Bad Firms", *Review of Industrial Organization*, Vol. 21, Number 3:305-324, 2002.

Bland, Eugene M., Black, Gregory S., Lawrimore, Kay, "Risk-reducing and Risk-enhancing Factors Impacting Online Auction Outcomes: Empirical Evidence from eBay Auctions", *Journal of Electronic Commerce Research*, Vol. 8, No. 4:236-243, 2007.

Borenstein, Severin & Saloner, Garth, "Economics and Electronic Commerce", *Journal of Economic* Perspectives, Vol. 15, No. 1:3-12, 2001.

Brown, Jennifer & Morgan, John,"Reputation in Online Auctions: The Market for Trust", *California Management Review*, Vol. 49, No. 1:61-81, 2006.

Brown, Mark & Muchira, Rose, "Investigating the Relationship between Internet Privacy Cocerns and Online Purchase Behavior", *Journal of Electronic Commerce Research*, Vol. 5, No. 1:62-70, 2004.

Brynjolfsson, Erik & Smith, Michael D., "Consumer Decision Making at an Internet Shopbot: Brand Still Matters", *The Journal of Industrial Economics*, Vol. 49, No. 4:541-558, 2001.

Cabral, Luís M.B., "The Economics of Trust and Reputation: A Primer", *Technical Report*, http://pages.stern.nyu.edu/~lcabral/reputation/Reputation_June05.pdf, 2005.

Cabral, Luís M.B., Hortaçsu, Ali: "The Dynamics of Seller Reputation: Evidence from eBay", *NYU Working Paper*, No. 2541/26094, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1282525, 2006.

Cortina, Jose M., "Interaction, Nonlinearity and Multicollinearity: Implications for Multiple Regression", *Journal of Management*, Vol. 19, No. 4:915-922, 1993.

Dellarocas, Chrysanthos, "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms", *Management Science*, Vol. 49, No. 10:1407-1424, 2003.

Dewan, Sanjeev & Hsu, Vernon, "Adverse Selection in Electronic Markets: Evidence from Online Stamp Auctions", *The Journal of Industrial Economics*, Vol. 52, No. 4:497-516, 2004.

Dua, Wenjing, Gu, Bin & Whinston, Andrew B., "Informational Cascade and Software Adoption on the Internet: An Empirical Investigation", *MIS Quarterly*, Vol. 33, No. 1:23-48, 2009.

Friedman, Batya, Kahn, Peter H. Jr. & Howe, Daniel, C., "Trust Online", *Communications of the ACM,* Vol. 43, No. 12:36-40, 2000.

Fuller, Mark A., Serva, Mark A. & Benamati, John "Skip", "Seeing Is Believing: The Transitory Influence of Reputation Information on E-Commerce Trust and Decision Making", *Decision Sciences*, Vol. 38, No. 4:675-699, 2007.

Gan, Lydia, He, Shujia, Tingli, Huang & Tan, Jiebin, "A Comparative Analysis of Online Grocery Pricing in Singapore", *Electronic Commerce Research and Applications*, 6:474-483, 2007.

Grover, Varun, Jaejoo, Lim & Ayyagari, Ramakrishna, "The Dark Side of Information and Market Efficiency in E-Markets", *Decision Sciences*, Vol. 37, Issue 3:397-324, 2006.

Houser, Daniel & Wooders, John, "Reputation in Auctions: Theory and Evidence from eBay", *Journal of Economics & Management Strategy*, Vol. 15, Number 2:353-369, 2006.

Hörner, Johannes, "Reputation and Competition", *The American Economic Review*, Vol. 92, No. 3:644-663, 2002.

Klein, Benjamin & Leffler, Keith, B., "The Role of Market Forces in Assuring Contratctual Performance", *The Journal of Political Economy*, Vol. 89, No. 4:615-641, 1981.

Jacobson, David & O'Callaghan-Andréosso, Bernadette, Industrial Economics and Organization: A European Perspective, The McGraw-Hill Book Companies, London, 1996.

Kim, Dongmin & Benbasat, Izak, "Trust-Related Arguments in Internet Stores: A Framework for Evaluation", *The Journal of Electronic Commerce Research*, Vol. 4, No. 2:49-64, 2003.

Kim, Hee-Woong & Xu, Yunjie, "Drivers of Price Premium in E-markets", *Communications of the ACM*, Vol. 50, No. 11:91-95, 2007.

Koenker, Roger & Hallock, Kevin F., "Quantile Regression: An Introduction", *http://www.econ.uiuc.edu/~roger/research/intro/rq.pdf*, 2001.

Kotler, Philip & Keller, Kevin L., Marketing Management (12th Edition), Pearson Prentice Hall 2006.

Kumaraguru, Ponnugaram, Acquisiti, Alessandro & Cranor, Lorrie Faith, "Trust Modelling for Online Transactions", *Working Paper*, http://www.cs.cmu.edu/~ponguru/pk_aa_lc_pst_2006.pdf, 2006.

Leiter, Daniel B. & Warin, Thierry, "An Empirical Study of Price Dispersion in Homogeneous Goods Markets", *Middlebury College Economics Discussion Paper*, No. 07-10, http://sandcat.middlebury.edu/econ/repec/mdl/ancoec/0710.pdf,, 2007.

Lin, Zhangxi, Li, Dahui, Jamananchi, Balaji & Huang, Wayne, "Reputation Distribution and Consumer-to-Consumer Online Auction Market Structure: An Exploratory Study", *Decision Support Systems*, 41: 435-448, 2006.

Livingston, Jeffrey A., "How Valuable Is a Good Reputation? A Sample Selection Model of Internet Auctions", *The Review of Economics and Statistics*, 87(3): 453-465, 2005.

Lucking-Reiley, David, Bryan, Doug, Prasad, Naghi & Reeves, Daniel, "Pennies from eBay: The Determinants of Price in Online Auctions", *The Journal of Industrial Economics*, Vol. 55, No. 2:223-233, 2007.

Mailath, George J. & Samuelson, Larry, "Who Wants a Good Reputation?", *The Review of Economic Studies*, Vol. 68, No. 2:415-441, 2001.

McDonald, Cynthia G., Slawson, Carlos Jr., "Reputation in an Internet Auction Market", *Economic Inquiry*, Vol. 40, No. 3:633-650, 2002.

Melnik, Mikhail I. & Alm, James, "Does a Seller's eCommerce Reputation Matter? Evidence from eBay Auctions", *The Journal of Industrial Economics*, Vol. L, No. 3:337-349, 2002.

Milgrom, Paul & Roberts, John, "Predation, Reputation, and Entry Deterrence", *Journal of Economic Theory*, 27:280-312, 1982.

Odom, Marcus D., Kumar, Anand & Saunders, Larry, "Web Assurance Seals: How and Why They Influence Consumers' Decisions", *Journal of Information Systems*, Vol. 16, No. 2:231-250, 2002.

Pan, Xing, Ratchford, Brian T. & Shankar, Venkatesh, "Can Price Dispersion in Online Markets Be Explained by Differences in E-Tailer Service Quality", *Journal of the Academy of Marketing Science*, Vol. 30, No. 4:433-445, 2002.

Pan, Xing, Ratchford, Brian T. & Shankar, Venkatesh, "Price Dispersion on the Internet: A Review and Directions for Future Research", *Working Paper*, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=485863, 2003.

Pate, Jennifer, "Seller Reputation as a Determinant of Price in Online Auctions: Evidence from Gift Card Sales", *Working Paper*, http://myweb.lmu.edu/jpate/reputation.pdf, 2006.

Pavlou, Paul A., & Dimoka, Angelika, "The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums and Seller Differentiation", Vol. 17, No. 4:392-414, 2006.

Professional eBay Sellers Association, "Unhealthy Marketplace Dynamics – Seller Perspective", http://www.gopesa.org/news/index.cfm?page=ebayunheatlhymarketplace2007, 2007.

Rao, Akshay R. & Bergen, Mark E., "Price Premium Variations as a Consequence of Buyer's Lack of Information", *Journal of Consumer Research*, Vol. 19:412- 423, 1992.

Resnick, Paul & Zeckhauser, Richard, "Trust among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System" in *The Economics of Internet and E-Commerce* (ed. Michael R. Baye), Vol. 11 of Advances in Applied Microeconomics, Amsterdam, Elsevier Science, pp. 127-157, 2002.

Resnick, Paul, Zeckhauser, Richard, Friedman, Eric & Kuwabara, Ko, "Reputation Systems", *Communications of the ACM*, Vol. 43, No. 12:45-48, 2000.

Resnick, Paul, Zeckhauser, Richard, Swanson, John & Lockwood, Kate, "The Value of Reputation on eBay: A Controlled Experiment", *Experimental Economics*, 9:79-101, 2006.

Standifird, Stephen S, "Reputation and e-Commerce: eBay Auctions and the Asymmetrical Impact of Positive and Negative Ratings", *Journal of Management*, 27:279-295, 2001.

Sun, Chia-Hung, "Asymmetric Information and Returns to Reputation in Online Auctions", *The Manchester School*, Vol. 76. No. 3:336-348, 2008.

Sun, Chia-Hung & Hsu, Ming-Fei, "The Determinants of Price in Online Auctions: More Evidence from Quantile Regression", *University of Wollongong Economics Working Paper Series 2007*, WP 07-18, http://www.uow.edu.au/content/groups/public/@web/@commerce/@econ/documents/doc/uow039115.pdf, 2007.

Salop, Steven C & Stiglitz, Joseph, E., "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion", *The Review of Economic Studies*, Vol. 44, No. 3:493-510, 1977.

Shapiro, Carl, "Premiums for High Quality Products as Returns to Reputation", *The Quarterly Journal of Economics*, Vol. 98, No. 4:659-680, 1983.

Sharma, Subhash, Durand, Richard M., & Gur-Arie, Oded, "Identification and Analysis of Moderator Variables", *Journal of Marketing Research*, Vol. 18:291-300, 1981.

Smith, Michael D. & Brynjolfsson, Erik, "Consumer Decision-Making at an Internet Shopbot: Brand Still Matters", *The Journal of Industrial Economics*, Vol. 49, No. 4:541-558, 2001.

Stigler, G., "The Economics of Information", *Journal of Political Economy*, Vol. 69, No. 3:213-225, 1961.

Thirumalai, Sriram & Sinha, Kingshuk K., "Customization Strategies in Electronic Retailing: Implications of Customer Purchase Behavior", *Decision Sciences*, Vol. 40, No. 1:5-36, 2009.

Varian, Hal R., "A Model of Sales", *The American Economic Review*, Vol. 70, Issue 4:651-659, 1980.

Venkatram, N., "The Concept of Fit in Strategy Research: Towards Verbal and Statistical Correspodence", *Academy of Management Review*, Vol. 14, No. 3:423-444, 1989.

Williamson, Oliver, E., "Markets and Hierarchies: Some Elementary Considerations", *The American Economic Review*, Vol. 63, Issue 2:316-325, 1973.

Wilson, Robert, "Reputation in Games and Markets" in *Game-Theoretic Models of Bargaining*, Alvin E. Roth (ed.), pp. 27-62, Cambridge University Press, Cambridge, 1985.

Wooldrdige,Jeffrey M., Introductory Econometrics: A Modern Approach (2nd Edition), South-Western College Publishing, 2002.

Zhang, Jie, "The Roles of Players and Reputations: Evidence from eBay Online Auctions", *Decision Support Systems*, 42:1800-1818, 2006.

Zhou, Ming, Dresner, Martin & Windle, Robert, J., "Online Reputation Systems: Design and Strategic Practices", *Decision Support Systems*, 44:785-797, 2008.