Reputation Premium and Reputation Management: Evidence from the Largest e-Commerce Platform in China*

Ying Fan†
University of Michigan

Jiandong Ju‡
SHUFE,
Tsinghua University

Mo Xiao§
University of Arizona

January 26, 2016

Abstract

We study the life-cycle effects of reputation using a panel data set consisting of a 25% random sample of all sellers on China's largest e-commerce platform, Taobao.com. We find a substantial return to reputation, but only for established sellers. New sellers, in contrast, lower their prices to boost transaction volume and ratings. This reputation management by new sellers leads to a decrease in their revenue in the short run and even a decrease in their business' survival likelihood in the longer run. We show that such differential effects at different stages of a seller's business life-cycle can arise when the effect of reputation on future payoffs dominates that on current payoffs.

Keywords: reputation management, reputation dynamics, e-commerce

JEL: L14, L15, L81

*This paper was previously circulated under the title "Losing to win: reputation management of online sellers". We thank the Editor, two anonymous referees, Ginger Jin, Tobias Klein, Vineet Kumar, Philip Leslie, Gaston Llanes and participants in numerous seminars and conferences for their constructive comments. We also thank Emek Basker for her careful reading and feedback on an earlier draft. Mo Xiao acknowledges the McGuire Center for Entrepreneurship at the University of Arizona for financial support.

†Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109; yingfan@umich.edu.

‡School of International Business Administration, Shanghai University of Finance and Economics, Shanghai, China, and School of Economics and Management, Tsinghua University, Beijing, China; jujd@sem.tsinghua.edu.cn.

§Eller College of Management, University of Arizona, 1130 East Helen Street, Tucson, AZ, 85721; mxiao@eller.arizona.edu.

1 Introduction

Economists have long recognized that information asymmetries can lead to market inefficiencies. Uncertainty about product quality can hinder markets. At the extreme, it can potentially lead to market failure (Akerlof (1970)). Such information asymmetry problems may be particularly severe in e-commerce markets where buyers and sellers trade anonymously. To deal with these problems, almost all online platforms provide a reputation system that collects feedback from members regarding past transactions and makes them public by providing them online.

How does reputation affect the behavior and business performance of sellers? The intrinsic dynamic feature of reputation implies that the effect of reputation may be different for sellers at different stages of their business' life-cycle. For example, the theoretical literature on reputation dynamics (e.g., Shapiro (1983) and Holmstrom (1999)) shows that, in a repeated game with reputation, sellers may choose to realize lower or even negative profits initially in order to earn a reputation premium in the future. In other words, new sellers may behave differently and respond to changes in ratings differently from established sellers. While established sellers may enjoy the returns to reputation, new sellers may actively manage their reputation at the cost of short-run returns. Empirically, there is a large literature evaluating the (static) return to reputation,[1] but few studies have examined seller incentives to build up reputation for future gains.[2] In this paper, we bridge this gap by studying the dynamic effects of reputation, linking reputation management to reputation premia.

The data we use is a panel data set from China's largest e-commerce platform, Taobao.com. Taobao.com (henceforth, Taobao, which means "hunting for treasures" in Chinese) is an online retail platform where small businesses or individual entrepreneurs trade with prospective buyers. Founded by the Alibaba Group, Inc. in 2003, Taobao has become China's largest e-commerce platform. By the end of 2012, Taobao had close to 500 million registered users and more than 800 million product listings on any given day. It processed roughly $170 billion in transactions in 2012, more than Amazon and eBay combined (Economist (2013)).[3]

[1]For example, see Resnick and Zeckhauser (2002), Dellarocas (2003), Houser and Wooders (2006), Resnick, Zeckhauser, Swanson and Lockwood (2006), Cabral and Hortacsu (2010), Luca (2011), Anderson and Magruder (2012), Jolivet, Jullien and Postel-Vinary (forthcoming) and Cai, Jin, Liu and Zhou (2014).

[2]A notable exception is Mayzlin, Dover and Chevalier (2014), which studies hotels' manipulation of online reviews to deceive consumers. In another paper, Brown and Morgan (2006) provide a case study of an eBay user's reputation management: this seller sold jokes at one penny each in order to gain positive feedback, and then switched to selling undeveloped land after accumulating a high online reputation. We study how sellers manage reputation through "honest" behavior.

[3]Taobao data has been used by other researchers such as Chu and Manchanda (forthcoming), Ye, Xu, Kiang, Wu

On the Taobao site, a seller's reputation is highly visible to all parties of transactions. After each transaction, a buyer can rate a seller by leaving a positive (+1), neutral (0), or negative (-1) score. A seller's rating score is the cumulative sum of these feedback scores. The rating score is then categorized into grades, which are displayed most prominently on a seller's website and whenever a product of the seller is referenced or returned as a search result. This reputation system is similar to that of eBay (although Taobao uses price posting instead of auctions),[4] with one main difference. Taobao reports a user's rating score as a seller and that as a buyer separately,[5] whereas an eBay user's total rating score lumps her seller reputation and buyer reputation into a single statistic.[6] Note that due to the cumulative nature of Taobao's reputation system, what we refer to as reputation in the paper is not independent of seller "size" measured by the number of cumulative transactions.[7]

Our panel data consists of a 25% random sample of all sellers between March 2010 and April 2011 on Taobao, which covers more than a million Taobao sellers over 14 months. It provides information on both seller reputation and seller performance. For every month during our sample period, we observe each seller's revenue, transaction volume, and survival as a business. We also observe measures of her reputation as a seller. Using these data, our baseline empirical strategy is to regress these various outcome variables on reputation. We distinguish new sellers from established sellers in our regressions to assess potentially different reputation effects at different stages of a seller's business life-cycle.[8]

and Sun (2013) and Ju, Su and Xu (2013), the last of which is the closest to this paper. Ju, Su and Xu (2013) use a quantile regression threshold model to study how Taobao sellers' pricing strategies vary with reputation using trading data for the iPod Nano from Taobao.

[4] This relates our research to numerous studies on eBay reputation. For example, Houser and Wooders (2006), Cabral and Hortacsu (2010), Elfenbein, Fisman and McManus (2012), Klein, Lambertz and Stahl (2015), Saeedi (2014) and Einav, Kuchler, Levin and Sundaresan (2015).

[5] On both sites, a seller can also rate a buyer after each transaction.

[6] eBay does separately report some *seller* rating measures of a user such as "item as described", "communication", "shipping time" and "shipping and handling changes." However, the most prominently displayed rating is the general score, which depends on both seller ratings and buyer ratings of a user.

[7] Since such a "cumulative" reputation system is used by eBay, research papers using eBay data also use a similar terminology. For example, Bajari and Hortacsu (2004) in the section on "Reputation Mechanisms" write that "online auction sites rely on voluntary feedback mechanisms, in which buyers and sellers alike can post reviews of each others' performance [...] The most prominently displayed summary statistic, which accompanies every mention of a user ID on eBay's web pages, is the number of positives that a particular user has received from other unique users, minus the number of negatives."

[8] Luca (2011), in a robustness analysis, breaks down the effect of Yelp ratings on restaurant revenue by restaurant size. He finds that the effect of reputation on revenue is positive for restaurants of all sizes. By contrast, we find that the effect is negative for new sellers but positive for established sellers. Our explanation for the different findings is the following: while lowering prices and selling more typically lead to an increase in ratings on Taobao (due to the cumulative nature of Taobao's reputation system), this is not true for restaurant ratings on Yelp. As a result, restaurants do not manage their Yelp ratings, at least not by lowering prices. In fact, to validate the regression discontinuity design in the paper, Luca (2011) devotes a section to rule out the possibility of restaurants manipulating

To identify the causal effect of reputation, we exploit the panel structure of our data and incorporate seller fixed effects and month fixed effects to capture unobserved seller-specific heterogeneity. There might also be time-varying seller-specific unobservable shocks. When the seller/month-specific shocks are uncorrelated with our reputation measures, the fixed-effect regressions identify the causal effect of reputation that we are interested in. To allow for a more general case where the seller/month-specific unobservable shocks are correlated with a seller's rating (e.g., when the unobservable shocks affect a seller's continuation decision), we use an instrumental variable strategy. As will be explained later in detail, we instrument a seller's reputation measures with her cumulative transaction volume as a buyer, which we argue is correlated with the seller's rating measures, and can be reasonably assumed to be exogenous. The exogeneity relies on the assumption that the seller/month-specific shocks (time-varying shocks over and above the seller fixed effect) are serially independent. We argue that the bias would be against our main findings when such a serial independence assumption does not hold.

Our results show that seller reputation has a substantial immediate return, but only for established sellers. Established sellers are able to charge higher prices, sell larger volumes, and receive more revenue as their reputation ratings increase. In particular, we find that an increase in reputation by one rating grade leads to an increase in monthly revenue by 38%. As a result, established sellers with better rating grades are more likely to survive over time. For instance, an increase in reputation by one rating grade (which varies between 0 and 18 in our sample) leads to an increase in a seller's 12-month survival likelihood by about 9 percentage points (the average 12-month survival likelihood of established seller is 79%).

However, the results are different when we turn to new sellers. An increase in reputation by one rating grade leads to a decrease in the monthly revenue by about 17%. Further investigation suggests that when their reputation increases, these new sellers lower their prices in order to boost their transaction volume. In short, not only do new sellers on average charge lower prices and earn less revenue per month than established sellers, they also decrease their prices and receive less revenue as their ratings increase. These findings suggest that as new sellers' ratings increase, they have more incentives to cut their prices in order to gain better ratings in the future.

Note that the causal effect of reputation we estimate through our reduced-form analyses includes both the direct effect of reputation and the effect of reputation through affecting seller behavior. While we cannot fully separate out seller behavior, the reduced-form results nevertheless shed light

their ratings.

on it. This is because a key finding of the paper is that as reputation increases, monthly revenue *decreases* but transaction volume increases for new sellers. As long as reputation has a positive impact on demand, such an empirical finding – by the Envelope Theorem – rejects a static model where a seller maximizes only her current profit,[9] indicating that sellers behave dynamically. Our preferred explanation of this finding is: as new sellers' ratings increase, they have more incentives to cut their prices in order to boost transaction volume and hence gain better ratings in the future (perhaps because of an increasing marginal return to reputation). As a result, the overall revenue decreases.

While an estimated negative effect of reputation on revenue and a positive effect of reputation on transaction volume indicate that sellers dynamically manage reputation, seller reputation management does not necessarily imply that the overall effect of reputation on revenue is negative. To better understand when the negative effect of reputation on revenue can arise and to provide some intuition on reputation management and reputation dynamics, we provide a simple two-stage dynamic model in the appendix of the paper. In this model, a seller's rating in the second period depends directly on her rating and her sales in the first period. This is a common feature of many e-commerce platforms, including eBay.com and Taobao.com. Due to the dependency of the second-period rating on the first-period sales, the seller's pricing decision in the first period has both a static effect on her current revenue and a (negative) dynamic effect on her future revenue. When the effect of a seller's rating on demand is sufficiently large, her optimal first-period price (as well as revenue) may decrease with her first-period rating. This is because an increase in her first-period rating may increase the marginal benefit from lowering prices in order to gain better ratings in the next period. Her optimal second-period price and revenue, however, both increase with her second-period rating. In summary, our simple model illustrates why reputation management leads to different effects of reputation for sellers at different stages of their business' life-cycle. Our empirical findings suggest that the dynamic effect of reputation highlighted in the model dominates the static effect.

To check the robustness of our results, we have conduct a series of robustness analyses and a fixed-cohort analysis. One potential concern with the interpretation of our estimation results is that new and established sellers may be distinctively different. We control for their difference in level by including seller fixed effects in the regressions. However, to allow for the possibility that

[9]By the Envelope Theorem, as long as reputation has a positive impact on demand, a seller's profit is increasing in reputation. Since a seller's transaction volume – and hence variable costs – increase with her reputation (according to the empirical finding), her revenue must also increase with reputation, contradicting the key finding of the paper.

they are also different in slope (i.e., in how they react to higher ratings), we conduct a fixed-cohort analysis where we follow a fixed cohort of new sellers' activities for about a year to investigate how the effect of ratings changes when a seller transits from new to established. Consistent with the baseline results, these fixed-cohort analyses indicate that as new sellers become more established, they cut prices less, sell even larger volumes and earn even more monthly revenue as their ratings increase, confirming the life-cycle effect of reputation.

Finally, we also find different effects of reputation on sellers' survival odds. While established sellers with better rating grades are more likely to survive over time, better-reputed new sellers have lower survival odds. In other words, new sellers' reputation management activities not only lead to a negative correlation between reputation and revenue in the short run but also have adverse effects on their eventual survival likelihood. Note that this result on survival does not necessarily imply that such sellers are irrational or risk-loving. In fact, when the reputation premium that a seller can enjoy conditional on surviving dominates the adverse effect of a decreased survival likelihood, it is rational for the seller – even a seller who is not risk-loving – to aggressively manage her reputation at the cost of decreasing her survival likelihood.

To summarize, our results show that new sellers' growing reputations motivate active reputation management, leading to a negative correlation between reputation and revenue in the initial phase of a seller's business life-cycle. These findings are consistent with the incentives described in the existing theoretical literature on reputation dynamics. While reputation dynamics and its implications on market design are extensively studied in theory, there has been little empirical work documenting how reputation concerns affect a seller's behavior. Two examples in this literature are Chevalier and Ellison (1999) and Hong, Kubik and Solomon (2000), which study mutual fund managers' and security analysts' decisions, respectively. They find that younger managers and analysts behave strategically due to the incentives generated by career concerns. There has been even less work on how such incentives change in response to changing reputations. To the best of our knowledge, Cabral and Hortacsu (2010) is the only work that provides systematic empirical evidence along this direction. Cabral and Hortacsu (2010) construct a panel data set using eBay sellers' feedback histories as proxies for seller transaction histories, and find an increase in the rate of negative feedback after their first negative feedback. This result implies that a seller has less incentives to exert effort after getting their first negative feedback. In our panel data, we directly observe measures of seller actions and outcomes in addition to seller reputation. This enables us to provide direct evidence on how a seller's actions change in response to her changing reputation.

More importantly, we not only study how reputation affects seller behavior, but also how such effects change over the life-cycle of a seller. We find stark differences in how reputation affects actions and outcomes of new and established sellers, illustrating the dynamic effects of reputation. Our work links reputation management with reputation premia, a linkage emphasized by the theoretical literature on reputation but sparsely documented in empirical work.

The rest of the paper is organized as follows. Section 2 describes Taobao and the data we use. Section 3 develops our empirical framework. Section 4 presents the results and discusses their implications. Section 5 offers concluding remarks. The Appendix provides first-stage results for regressions using instrumental variables, robustness analyses, and a simple theoretical model that explains our empirical findings.

2 Taobao and the Data

2.1 Taobao and its Online Feedback System

E-commerce in China is a fast-growing, multi-billion dollar market.[10] In this market, Taobao is the undisputed market leader. Taobao was launched by the Alibaba Group, Inc. on May 10, 2003 to provide a platform for individuals or small businesses to trade with customers in an online marketplace. With an innovative website service and technical support at little cost to sellers,[11] Taobao soon dominated all other Chinese e-retailers, including eBay China and Amazon.cn.[12] As of January 2010 (two months before our data were collected), Taobao had approximately 180 million registered users, of which about 2 million were sellers. Taobao's 2010 transactions totaled approximately $60 billion. Taobao is still growing fast. By the end of 2012, it had close to 500 million registered users and more than 800 million product listings on any given day.[13] As mentioned, Taobao processed about $170 billion in transactions in 2012, more than Amazon and eBay combined.

In addition to offering a marketplace for buyers and sellers to meet, Taobao also provides an escrow payment service (called "Alipay") to facilitate transactions. The escrow payment service

[10] China's e-commerce market was the second largest in the world (after the U.S.) in 2011, with a total sales of $120 billion. Its compounded growth rate of 120% between 2003 and 2011. This is compared to a compound growth rate of about 20% in Germany, South Korea, U.S. and U.K. (see Dobbs, Chen, Orr, Manyika, Chui and Chang (2013)).

[11] Advertising has been Taobao's main income source. Two highlights of Taobao's services are Aliwangwang, an online instant messaging system to facilitate buyer-seller communication, and Alipay, an escrow payment service described later.

[12] eBay eventually shut down its site in China in 2006.

[13] http://www.taobao.com/about/intro.php?spm=0.0.0.0.RgqosI, accessed in July 2013.

requests buyers to deposit money with Alipay at the time of purchase to ensure that buyers have enough funds to pay. It also allows buyers to verify their purchase before instructing Alipay to release funds to sellers. When a buyer is dissatisfied with her purchase (quality defects, delayed delivery, etc.), a buyer can request Taobao to withhold her payment until the issue is solved. If the buyer and the seller cannot reach an agreement, Taobao will moderate and eventually issue a ruling for unresolved disputes.

To prevent sellers from registering multiple accounts and changing their online identity strategically, Taobao requires that all sellers register their accounts with valid national identification cards. A national identification card in China contains information on a citizen's full name, gender, ethnicity, date of birth, address, identification number and a photo of the individual. To register an account, a seller needs to upload a scan of her national identification card as well as a picture of herself holding her national identification card. The latter requirement makes it hard for an individual to open an account using others' identification cards. Overall, by linking a seller's online identity to her personal identity, Taobao makes it difficult for a seller to have multiple accounts or to discard an old account and reappear with a new account.

Besides requiring sellers to register an account with a valid national identification card, Taobao employs a feedback system to build trust among participants in transactions. On Taobao, a buyer can rate a seller after each transaction. When a buyer purchases multiple commodities in one order, each traded commodity is considered a transaction.[14] The default rating score is positive (+1) unless it is overwritten with a neutral (+0) or negative (-1) score.[15] A seller's rating score is the cumulative sum of these feedback scores from each transaction.[16] Given the cumulative nature of the rating score, a seller's rating score depends on her cumulative transaction volume, a feature captured by our theoretical model in the appendix and shared by other e-commerce platforms such as eBay. A seller's rating score is then categorized into one of 20 grades, each represented by a

[14] When a buyer purchases multiple units of one commodity in an order, however, this is considered only one transaction.

[15] The buyer has 15 days to submit her feedback after a transaction is completed. If she does not submit feedback on time, Taobao will assign a positive rating to the seller on her behalf. Note that a transaction is considered completed only after the buyer confirms the delivery and accepts the delivered products, or after a certain amount of time since the shipment, depending on the speed of the delivery, if the buyer does not dispute the transaction. In other words, a transaction with a pending dispute is not considered completed. With this mechanism, even if no buyer chooses to rate the seller and all ratings are positive by default, the rating score is still a meaningful measure of seller reputation.

[16] To make sure that sellers cannot manipulate their ratings through fake transactions (for example, asking friends to "purchase" their products repeatedly and give positive feedback for each transaction), Taobao limits a given buyer to six positive comments per month in its calculation of a seller's rating score. In other words, for any buyer/seller pair, the buyer can contribute to at most 6 positive rating scores for the seller per month. Note that this rule does not invalidate the positive correlation between a seller's transaction volume and her rating score.

system of hearts, diamonds, crowns and golden crowns. See Table 1 for the mapping from the rating score to the rating grade on Taobao. These twenty grades are well recognized by Taobao users. For example, a "crown" seller is immediately considered a highly-reputed and successful seller. These rating grades are displayed prominently on a seller's shop sites as well as when sellers or their products appear in search results.[17]

Table 1: Seller Rating Score and Grade on Taobao

Seller Rating Score	Seller Rating Grade
< 4	
4 – 10	❤
11 – 40	❤❤
41 – 90	❤❤❤
91 – 150	❤❤❤❤
151 – 250	❤❤❤❤❤
251 – 500	💎
501 – 1,000	💎💎
1,001 – 2,000	💎💎💎
2,001 – 5,000	💎💎💎💎
5,001 – 10,000	💎💎💎💎💎
10,001 – 20,000	👑
20,001 – 50,000	👑👑
50,001 – 100,000	👑👑👑
100,001 – 200,000	👑👑👑👑
200,001 – 500,000	👑👑👑👑👑
500,001 – 1,000,000	👑
1,000,001 – 2,000,000	👑👑
2,000,001 – 5,000,000	👑👑👑
5,000,001 – 10,000,000	👑👑👑👑
10,000,001 –	👑👑👑👑👑

On Taobao, a seller can also rate a buyer after each transaction. Taobao distinguishes a registered user's rating score as a seller from her rating score as a buyer. In other words, a seller's seller rating score on Taobao is based on the feedback she gets as a seller only. Thus, a seller cannot manipulate her rating score by increasing her purchasing activity.

[17]Taobao also reports other rating measures such as "item as described", "customer service" and "shipping and delivery". But the overall rating is the most prominently displayed rating measure.

2.2 Data

Our proprietary data consist of a 25% random sample of all sellers on Taobao between March 2010 and April 2011. To draw this random sample, Taobao collected the universe of sellers who have sold at least one item during this time, took a 25% random sample of these sellers, and provided us data on them during these 14 months. On the site, Taobao distinguishes between Marketplace sellers and Mall sellers. Marketplace sellers are mostly individuals or small businesses. By contrast, Mall sellers are companies who are either brand owners or authorized distributors. Our 25% random sample includes 1,322,546 Taobao Marketplace sellers and 9,462 Taobao Mall sellers. Since Mall sellers also have an offline reputation, we drop them from our sample. To exclude those who are occasional sellers only (for example, those who try to resell their personal items), we also drop Marketplace sellers who are inactive in one third of the time span between their first and last appearances in the data, which amounts to about 18.61% of the initial sample of Marketplace sellers and about 2.53% of the total revenue and 2.51% of the total transactions observed in the data by these Marketplace sellers.[18] Finally, we drop sellers with obvious data reporting errors.[19] In the end, we are left with 1,063,167 unique sellers.

For each seller in our sample, we observe her gender, age, birth province, residence province and residence city. An average Taobao seller in our sample is 29 years old, 54% are women, and 37% have immigrated from their birth province to their residence province.

For each month that a seller is in the data, we observe her revenue,[20] transaction volume (i.e., the number of transactions: as explained in Section 2.1, on Taobao, each item in an order is considered a transaction), and main business category in the current month. We also observe her cumulative number of selling and buying transactions since her registration on Taobao. Finally, we observe several measures of seller reputation: rating score, rating grade, and percentage of positive ratings.

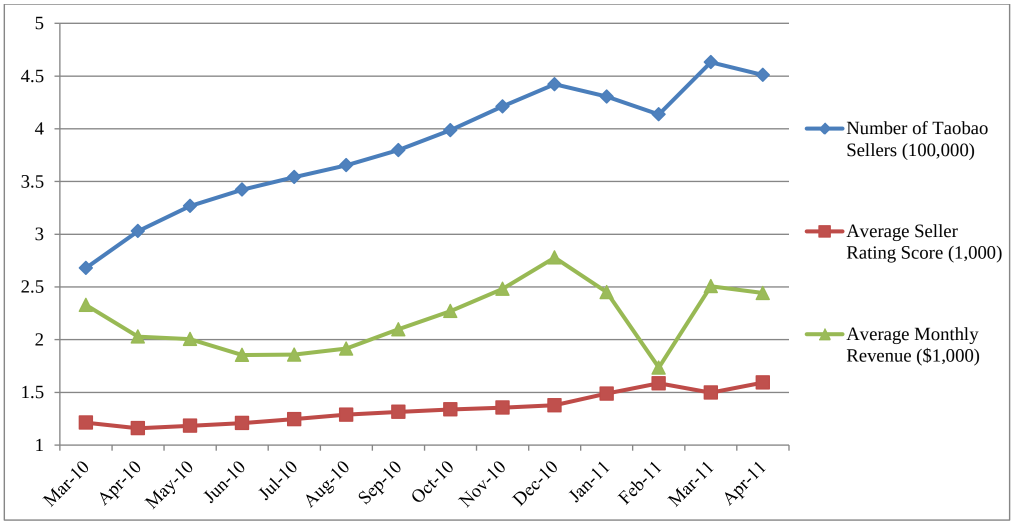
Figure 1 shows the time series for the number of Taobao sellers in our sample, their average rating score and average monthly revenue in the 14 months of our sample period. From Figure 1, we see that the number of sellers rapidly increases over time, from a little over 260,000 in March 2010 to about 450,000 in April 2011. We also see that the average seller rating score and average

[18] Our results are robust when we drop sellers who are inactive for one half of the time span.

[19] An example of a data reporting error would be if the number of *cumulative* seller transactions decreases over time.

[20] We convert revenue in RMB to revenue in U.S. dollars using the exchange rate in July 2011 (1 U.S. dollar equals 6.472 RMB).

Figure 1: The Evolution of Taobao Sellers in our Sample



monthly revenue also increase slightly over time. Even though new sellers join Taobao every month, potentially bringing down the average rating score, the growth in the rating scores of the existing and surviving sellers offsets this effect. Note that, in February 2011, both the number of Taobao sellers and their average monthly revenue took a plunge. This is because the Chinese New Year fell into this month and the entire nation – especially the post office and various parcel delivery companies, which deliver most Taobao transactions – was on a break. While a large number of sellers just take a break during the holidays, some consider the time a natural point for exiting the market. As a result, we see a sharp drop in the number of Taobao sellers and the average monthly revenue in this month.

Table 2 provides the distribution of seller ratings in our random sample. For each rating grade symbol, we assign a numerical rating grade. In our sample, the highest grade reached is 18, or, equivalently, 3 golden crowns. About 40% of seller-months achieve the diamond status, while fewer than 2.5% achieve the crown status.

3 Empirical Framework

3.1 Defining New and Established Sellers

In order to study the potentially different effects of seller reputation on a seller's behavior and business performance at different stages of a seller's business life-cycle, we distinguish new sellers

Table 2: The Distribution of Seller Ratings

Seller Rating Score	Seller Rating Grade[a]	Frequency	Percent	Cumulative
Below 4	0	393,803	7.35	7.35
4 – 10	1	441,247	8.23	15.58
11 – 41	2	867,299	16.18	31.76
41 – 90	3	632,231	11.8	43.56
91 – 150	4	424,077	7.91	51.47
151 – 250	5	419,987	7.84	59.31
251 – 500	6	639,662	11.93	71.24
501 – 1,000	7	535,426	9.99	81.23
1,001 – 2,000	8	404,274	7.54	88.77
2,001 – 5,000	9	338,159	6.31	95.08
5,001 – 10,000	10	135,936	2.54	97.62
10,001 – 20,000	11	74,895	1.40	99.02
20,001 – 50,000	12	39,300	0.73	99.75
50,001 – 100,000	13	8,819	0.16	99.91
100,001 – 200,000	14	2,954	0.06	99.97
200,001 – 500,000	15	1,292	0.02	99.99
500,001 – 1,000,000	16	236	4.40e-05	100
1,000,001 – 2,000,000	17	101	1.89e-05	100
2,000,001 – 5,000,000	18	30	5.60e-06	100
Seller/months		5,359,728		

[a]In our 25% random sample, the highest grade reached is 18 (3 golden crowns).

from established sellers for our empirical analyses. In general, a new seller in our paper is defined as a seller who has started selling recently, and who is yet to reach a certain level of transaction volume; and an established seller is one who has been selling on Taobao for a while and has reached a certain level of transaction volume. Specifically, we define new sellers and established sellers based on when they first appear in our sample and their cumulative transaction volume at this first appearance. We require a seller to satisfy both criteria to be considered a new or an established seller. Table 3 specifies the cutoff points for our baseline definitions.

Table 3: Baseline Definition of New Sellers and Established Sellers

	First appears in the data	Cumulative transactions at the first appearance
New Seller	later than Month 1	≤ 30
Established Seller	in Month 1	> 250

We include the first month criterion (when a seller first appears in the data) in defining new and established sellers because a seller who is not in our data in Month 1 is more likely to have just started selling. Conversely, a seller who is in our data in Month 1 is more likely to have been selling on Taobao for a while.[21] In our baseline definition, an established seller first appears in Month 1 and has completed more than 250 transactions by the end of Month 1. We choose 250 as the cutoff because it takes a score of 251 for a seller to reach the well-known milestone diamond status. Similarly, by this definition, a new seller first appears later than Month 1 and has not completed more than 30 transactions by the end of her first appearance month. We choose 30 as our new-seller cutoff because for sellers who first appear later than Month 1, the majority (75%) have not completed more than 30 transactions by the end of her first appearance month. Doing so allows us to err on the side of caution and exclude sellers who have sold a lot by their first appearance month from new sellers. Such sellers may just have taken a break and returned to business in our 14-month window.

This is our baseline definition. In robustness analyses, we vary the starting month criterion and the transaction volume criterion, and show that our results are robust to these variations. Our results are also robust to omitting the starting month criterion as part of the definition. We show these robustness analyses in Appendix B. Note that these definitions are seller-specific and define two exclusive groups of sellers in our data. In Section 4.2, we also conduct analyses focusing on a fixed cohort of sellers to study the effect of reputation before and after they reach a certain rating grade. In these fixed-cohort analyses, the concepts of "new" and "established" are dynamic, i.e., seller/month-specific. In other words, a new seller can become an established one after accumulating a certain level of reputation.

In summary, in defining new sellers and established sellers, we use a conservative baseline specification and subsequently conduct various robustness analyses using different definitions. Being conservative in defining new sellers and established sellers means that some sellers are neither new nor established and are thus dropped from our sample. Our final sample includes 329,169 new sellers corresponding to 1,349,124 seller/month observations and 104,138 established sellers corresponding to 1,234,176 seller/month observations.[22] Table 4 reports the summary statistics for these new sellers and established sellers. These statistics show that new sellers, on average, have much lower rating scores. They also have much lower average monthly revenue and average transaction volume. The average percentage of positive ratings is about the same for new sellers and established sellers,

[21]Our data does not provide information on when a seller starts her business.

[22]As will be explained in Section 3.2, we use lagged rating measures as our explanatory variable in the empirical analysis, which means that we lose one month of observations.

although the latter group exhibits a much smaller variance.

Table 4: Summary Statistics: New Sellers vs. Established Sellers

	New Sellers		Established Sellers	
	Mean	Std. Dev.	Mean	Std. Dev.
Monthly Revenue ($)	876.613	8459.181	5966.275	44993.89
Monthly Transactions	40.237	188.973	327.73	2386.051
Lagged Seller Rating Score	109.068	422.643	4670.873	23347.87
Lagged Seller Rating Grade	2.461	1.912	8.196	1.609
Lagged % Seller Positive Ratings (%)	99.496	2.682	99.467	0.800
Observations (Seller/months)	1,349,124		1,234,176	
Sellers	329,169		104,138	

3.2 Empirical Specification

To identify the heterogeneous effect of reputation on outcomes for new and established sellers, we run the following regression separately for these two groups:[23]

$$\begin{aligned} y_{it} &= \alpha_1(RatingGrade)_{i,t-1} + \alpha_2(RatingScore)_{i,t-1} + \alpha_3(\%PositiveRating)_{i,t-1} \\ &\quad + \mu_i + \omega_t + \varepsilon_{it}. \end{aligned} \tag{1}$$

In this equation, we index a seller by i and a month by t. We estimate this equation for various dependent variables y_{it}, such as seller i's revenue in month t or her transaction volume in month t. These outcome variables are aggregated over a month, whereas only a snapshot of a seller's reputation profile on the 15^{th} of each month is reported in the data. We therefore use lagged reputation measures as our independent variables. Using lagged reputation variables also alleviates the reverse causality concern because current sales do not affect the previous month's rating score. We include three measures of seller reputation available in the data. $(RatingGrade)_{it-1}$ is an integer from 0 to 18. $(RatingScore)_{it-1}$ is the rating score. $(\%PositiveRating)_{it-1}$ is the percentage of positive ratings. We include a seller fixed effect μ_i in our estimation to control for unobserved seller heterogeneity. We also include a month dummy ω_t to capture any seasonality effects or macroeconomic shocks. Lastly, ϵ_{it} is an error term to capture seller/month-specific unobserved heterogeneity, which we assume to be i.i.d. across sellers and over time.

[23] Doing so is equivalent to using data on all sellers and estimating a regression where we interact all time-varying variables with a new-seller dummy variable and an established-seller dummy variable.

In the above empirical specification (1), we are able to include the seller and month fixed effects because of the panel nature of our data set. Including them, in particular the seller fixed effects, alleviates endogeneity concerns about our rating measures, especially the concerns over unobserved seller heterogeneity. For example, factors such as a well-organized website design or an easy-to-remember business name can affect a seller's rating score as well as her revenue. Such concerns have been an issue for previous research that uses only cross-sectional variation in seller reputation.[24]

Using seller fixed effects, we argue, alleviates most of the endogeneity concerns in studying the effect of reputation. This is because we are effectively using within-seller variation in seller ratings and seller outcome variables for identification, holding unobserved seller-specific attributes constant. We further use an instrumental variable strategy to deal with potential endogeneity problems arising when the seller-month specific error term ϵ_{it} is still correlated with a seller's previous month rating. For example, there might be a selection bias. In each month, we only observe sellers who choose to continue operating in this month. If both the shock ϵ_{it} and the seller's rating in the previous month affect her continuation decision, the error term and the seller's rating in the previous month can be correlated in our sample. For example, if both the revenue shock and seller ratings positively affect a seller's survival, then they are negatively correlated in our sample. Such a selection issue is similar to that highlighted in Olley and Pakes (1996), where they observe only surviving firms whose survival is influenced by both the unobserved productivity shock and their capital level.

The instrumental variables we use are constructed based on a seller's cumulative transaction volume as a buyer in the previous month.[25] We obtained a seller's cumulative transaction volume as a buyer in each month from Taobao, while a Taobao buyer does not observe this metric.[26] For this IV strategy to work, we need the last-period cumulative buyer transactions to be correlated with last-period seller reputation (relevance), to be uncorrelated with the current shock conditional on fixed effects (exogeneity) and to not affect the current-period dependent variable directly

[24] Jolivet, Jullien and Postel-Vinary (forthcoming) and Cai, Jin, Liu and Zhou (2014) are two exceptions. The former uses a panel of sellers on a French retail platform. The latter uses panel data from another Chinese e-commerce platform, Eachnet.com, and includes a seller cohort fixed effect in their regressions. In another study, Cabral and Hortacsu (2010) construct a panel using feedback histories as proxies for seller transaction histories. Finally, to address endogeneity concerns, Luca (2011) and Anderson and Magruder (2012) exploit the special institutional setting of Yelp.com and use a regression discontinuity design.

[25] Alternatively, we could use lagged variables as instrumental variables, for example, $(RatingScore)_{it-2}$ and $(Transactions)_{it-2}$. Doing so yields similar results. But we lose more than 10% of our sample by using lagged variables as instrumental variables.

[26] A buyer can imperfectly infer a seller's current cumulative buyer transaction volume from her ratings as a buyer (only if the buyer is willing to investigate a seller's detailed rating profile, which is a few clicks away from the seller's main shop site).

(exclusion).

A seller's cumulative transaction volume as a buyer is likely to be correlated with our endogenous variables, i.e., her seller ratings. We think the correlation mainly comes from time use on Taobao. The time that a seller spends on Taobao may influence both her buyer transaction volume as well as her seller transaction volume, hence her seller rating. For example, a seller who spends 16 hours answering customer inquiries on Taobao may also make a lot of her personal purchases on Taobao. This correlation is confirmed by the first-stage results in Table A.1 in Appendix A.

As for exogeneity, we need the instrumental variables to be weakly exogenous, i.e., the lagged cumulative buyer transaction volume is uncorrelated with the current shock. This condition is likely to hold if the shock is serially independent.[27] Given that we have included seller fixed effects to capture the persistent part of unobservable seller heterogeneity, we think the assumption that the shock (over and above the seller fixed effect) is independent over time is plausible.[28]

Finally, for the exclusion restriction, we need a seller's cumulative transaction volume as a buyer in the previous month to not affect the current-period dependent variable directly, through either the demand or the supply channel. It is unlikely that a seller's cumulative transaction volume as a buyer shifts demand because potential buyers on Taobao are unlikely to use the information on a seller's buyer volume – in addition to the seller's ratings as a seller – in making purchasing decisions.[29] It is also unlikely that a seller's buyer volume in the previous month affects the current-period supply. Such a supply channel could be: a seller buys inputs from Taobao and thus a larger amount of input purchase last month may lead to larger sales this month. But note that Taobao is a retail platform rather than a wholesale platform. Most Taobao sellers buy for personal use rather than to purchase inputs.

[27]This condition is also likely to hold if our instrumental variables satisfy contemporaneous exogeneity. If the lagged cumulative buyer transaction volume is uncorrelated with the lagged shock, it is likely to be uncorrelated with the current shock as well.

[28]When our instrumental variables do not satisfy the exogeneity condition, our estimates would be biased. One important result in our paper is that monthly revenue decreases with ratings for new sellers. If our IV was positively correlated with the error term, our IV results would be biased upwards. In other words, the bias would be against this finding. We therefore need to worry about the case where our IV is negatively correlated with the error term. There are two such scenarios: (1) the error term, i.e., a revenue shock is negatively serially correlated and the lagged cumulative buyer transactions is positively correlated with the lagged revenue shock; or (2) the shock is positively serially correlated and the lagged cumulative buyer transactions is negatively correlated with the lagged shock. We think scenario (1) is implausible because the revenue shock is unlikely to be negatively correlated over time especially given that we have controlled for monthly dummies and thus possible seasonality. Moreover, the second part of scenario (2) seems difficult to square with our first-stage regression results, which suggest that a seller's cumulative buyer transaction volume and her cumulative seller transaction volume are positively correlated. It is therefore unlikely that a seller's cumulative buyer transaction volume is negatively correlated with the revenue shock.

[29]Taobao's escrow payment service ("Alipay") requests buyers to deposit money with Alipay at the time of purchase to ensure that buyers have enough funds to pay. In other words, a user's transaction volume as a buyer does not reflect her reliability to make payments or other intrinsic characteristics of this user.

Given the above arguments on relevance, exogeneity and exclusivity, we use a seller's transaction volume as a buyer in the previous month to instrument for her seller rating measures in that month. Note that the seller rating grade in equation (1) is a function of a seller's rating score. We construct a grade variable from a seller's cumulative transaction volume as a buyer analogously and include it as an instrumental variable. Specifically, we create a grade variable based on the cutoff points in Table 1.

Table 5 presents the summary statistics for our instrumental variables while Table A.1 in Appendix A presents the first-stage results. The frist-stage results show that measures of a seller's transaction activities as a buyer are significant determinants of her seller reputation measures. The F-statistics in the first-stage regressions are all far above 10 and thus we can reject the null that the excluded instruments are irrelevant in the first-stage regressions. In another set of regressions, we find that there is a positive correlation between a seller's buyer transaction volume and her seller transaction volume even when we control for how long a seller has been selling.[30]

Table 5: Summary Statistics: Sellers' Cumulative Transaction Volume as a Buyer

	New Sellers		Established Sellers	
	Mean	Std. Dev.	Mean	Std. Dev.
Lagged Cumulative Buyer Transactions	77.280	247.531	202.537	337.984
Lagged Cumulative Buyer Transaction Grade	2.253	1.838	3.931	1.931
Observations (Seller/months)	1,349,124		1,234,176	
Sellers	329,169		104,138	

For the percentage of positive ratings, we could use the percentage of positive ratings given to the seller after a buying transaction. There is, however, little variation in this percentage as few sellers give buyers a neutral or negative feedback. Since there is no good instrument for the percentage of positive feedback, this precludes a causal interpretation of the coefficient of the percentage of positive seller ratings.[31]

[30] In such regression, we add dummy variables constructed based on months since the first time that a seller appears in the data as covariates. We run such regressions for new sellers only as all established sellers appear in the data in the first month. Therefore, the time of the firs appearance is not a good approximation for their entry time.

[31] Note that under the assumption that our instrumental variables are uncorrelated with the percentage of positive seller ratings, our estimates of the other coefficients are unbiased in the IV regressions. See Ackerberg, Crawford and Hahn (2011) for a proof. In our data, the correlation between the two instrumental variables and uninstrumented variable are, respectively, 0.027 and 0.018.

4 Results

4.1 The Effect of Ratings on Revenue and Transaction Volume

Using the empirical specification in the previous section, we now quantify the effect of reputation on a seller's behavior and business performance. Table 6 reports the results when we regress log monthly revenue on various rating measures.[32] We present three specifications, gradually adding more seller rating measures, and alternate between the OLS and IV results. We include seller fixed effects and month dummies in all specifications.

Table 6(a) presents the results for established sellers. As we add more reputation measures in the regression, we can see that coefficients remain relatively stable. The comparison of the OLS results and the IV results indicates a negative bias in the OLS regressions. This bias leads to an underestimation of the percentage change in a seller's monthly revenue by about 5% when her rating grade increases by 1. This bias could be generated by a negative correlation between the error term and ratings, which in turn can arise if both the error term and ratings are positively correlated with survival.

The last column of Table 6(a) shows the IV estimates when all three rating measures are included. Here, we see that a one point increase in the lagged rating score leads to very little gain. However, a one rating grade increase leads to a 38% increase in monthly revenue. This difference is not surprising because the rating grade is the most prominently displayed rating measure for sellers. Overall, our results for established sellers indicate a reputation premium for them: better reputation leads to higher monthly revenues.

Table 6(b) for new sellers, however, shows a different pattern. A higher rating does not seem to be associated with higher revenue for new sellers. In fact, the estimates suggest the opposite. The IV estimates in the last column of Table 6(b) indicate that a one rating grade increase leads to a roughly 17% decrease in monthly revenues.

To understand this result, we decompose monthly revenue into the number of transactions and the average revenue per transaction. As explained in Section 2.1, multiple items in one order are considered multiple transactions on Taobao. Thus, transaction volume measured on Taobao is closer to the concept of "quantity" than the number of completed orders. For the same reason, we consider the average revenue per transaction to be a rough measure of price. We then regress

[32] To avoid taking the logarithm of zero for seller/month observations with no sales, we add 1 Chinese cent (1/6.427 US cent) to revenue.

Table 6: The Effect of Reputation on log(Revenue)

(a) Established Sellers

	OLS	IV	OLS	IV	OLS	IV
L. Rating Grade	0.318***	0.427***	0.314***	0.354***	0.324***	0.378***
	(0.007)	(0.035)	(0.007)	(0.039)	(0.007)	(0.039)
L. Rating Score			2.228e-6***	4.640e-5***	2.455e-6***	4.230e-5***
			(4.518e-7)	(7.750e-6)	(5.085e-7)	(7.640e-6)
L. % Positive Ratings					27.396***	33.695***
					(2.616)	(1.580)
Seller fixed effects	yes	yes	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes	yes	yes
Within R-square	0.022		0.022		0.022	
Obs. (Seller/months)				1,234,176		
Sellers				104,138		

***: significant at the 1 percent level.

(b) New Sellers

	OLS	IV	OLS	IV	OLS	IV
L. Rating Grade	-0.079***	-0.287***	-0.091***	-0.167***	-0.092***	-0.166***
	(0.004)	(0.011)	(0.005)	(0.038)	(0.005)	(0.038)
L. Rating Score			1.030e-4***	-1.132e-3***	1.057e-4***	-1.152e-3***
			(1.291e-5)	(3.396e-4)	(1.287e-5)	(3.386e-4)
L. % Positive Ratings					2.194***	1.858***
					(0.216)	(0.215)
Seller fixed effects	yes	yes	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes	yes	yes
Within R-square	0.028		0.028		0.029	
Obs. (Seller/months)				1,349,124		
Sellers				329,169		

***: significant at the 1 percent level.

these two measures on seller ratings, using the same specification as in equation (1). Table 7 presents the IV regression results.[33] These results show that for established sellers, a higher rating grade contributes to a higher transaction volume and a higher average revenue per transaction. Specifically, an increase in the rating grade by one grade leads to an increase in the number of transactions by 26% and the average revenue per transaction by 4%. Both these effects contribute to the substantial return to reputation for established sellers. By contrast, for new sellers, a

[33] The dependent variables are log(1+transactions) and log(average revenue per transaction). We add 1 to transactions in order to avoid taking the logarithm of zero for seller/months observations with no sales. We drop such observations in the regression of average revenue per transaction. Consequently, the sum of the coefficients of these two regressions is close to but not the same as that of the revenue regression.

Table 7: The Effect of Reputation on Transactions and the Average Revenue per Transaction

	Established Sellers		New Sellers	
	log(1+Trans.)	log(Avg Revenue per Trans.)	log(1+Trans.)	log(Avg Revenue per Trans.)
L. Rating Grade	0.263***	0.036***	0.053***	-0.031***
	(0.016)	(0.010)	(0.015)	(0.011)
L. Rating Score	1.577e-5***	2.332e-6	-2.672e-3***	1.330e-3***
	(3.204e-6)	(2.032e-6)	(1.349e-4)	(1.024e-4)
L. % Positive Ratings	23.410***	1.230***	7.138e-4	0.717***
	(0.662)	(0.425)	(0.086)	(0.067)
Seller fixed effects	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes
Obs. (Seller/months)	1,234,176	1,189,225	1,349,124	1,233,678
Sellers	104,138	104,138	329,169	329,169

Note: This table reports the IV regression results. We add 1 to transactions to avoid taking logarithm of zero for seller/months with no transaction. We drop such observations for the regression of log(average revenue per transaction). ***: significant at the 1 percent level.

higher rating grade is associated with a higher transaction volume but lower average revenue per transaction. A one rating grade increase leads to a 5% increase in the number of transactions, but a 3% decrease in the average revenue per transaction. These results suggest that as the rating grade increases, new sellers cut prices in order to boost transaction volume and ratings.[34]

In summary, these results indicate that the effect of reputation as measured by the rating grade is different for new sellers and established sellers. For an established seller, an increase in her rating grade leads to an increase in the price, the transaction volume and, in turn, her monthly revenue. For a new seller, however, an increase in her rating grade leads to a decrease in price and revenue, but an increase in quantity. All these findings are consistent with the intuition explained in Appendix C, where we provide a simple two-period theoretical model to explain the tradeoff between the dynamic effect of reputation and its static effect. Under the dynamic consideration, a forward-looking seller has incentives to lower prices in the first period, which leads to higher ratings and hence higher payoffs in the second period. When the effect of a seller's rating on demand is sufficiently large, her optimal first-period price (and revenue) may decrease with her first-period rating. Our empirical findings suggest that for new sellers, the dynamic effect of reputation seems to dominate its static effect. New sellers, anticipating the long-run returns to higher reputation,

[34]The decreased average revenue per transaction could also be consistent with sellers switching their business category to a cheaper one. For example, a seller may switch from selling computers to selling computer accessories. However, restricting our sample to sellers who have never switched their main business category yields similar results.

engage in price promotions in order to reach for a higher rating grade. On the way to "race for diamonds", the aforementioned common practice among Taobao users, new sellers on average forgo the short-run return to reputation.

Appendix C also explains why the coefficient of the rating score may have a different sign from the coefficient of the rating grade. As mentioned, whether the dynamic effect of a rating measure dominates its static effect depends on how large the effect of the rating measure on demand is. Since the rating grade represented by a system of hearts, diamonds, crowns and golden crowns is more salient than the rating score, it is likely to have a larger effect on demand than the rating score. As a result, it is possible that while within a rating grade, a seller charges a higher price as her rating score increases (i.e., the static effect dominates), the seller charges a lower price as her rating grade increases (i.e., the dynamic effect dominates). This also explains the different effects of the rating grade and the rating score on the monthly transactions.

4.2 Fixed-Cohort Analysis

Our findings indicate that established sellers enjoy a reputation premium, but new sellers engage in aggressive reputation management activities and more so as their reputation increases, which leads to a negative correlation between reputation and revenue. However, one potential concern with this interpretation of our estimation results is that new and established sellers may be distinctively different. For example, established sellers might be different from sellers who started only after Taobao became well-known in China. In our empirical design, we control for most seller differences by including seller fixed effects in the regressions. However, these two groups of sellers may also differ in how they react to higher ratings. In other words, the different estimates of the effect of ratings on revenue, transactions and the average revenue per transaction across these two groups may not be explained by the heterogeneous effect of reputation for sellers at different stages of their business' life-cycle. Here, we address this concern by focusing on one fixed cohort of sellers to study how the effect of ratings changes when a seller transits from new to established.

Specifically, we focus on a subgroup of new sellers: new sellers who first appear in our data in Month 2 (April 2010). There are 32,443 such sellers corresponding to 222,935 seller/months observations.[35] We next define a seller/month-specific dummy variable to indicate whether the seller is "established" in a given month. This variable has a value of 1 if the seller rating score

[35] There are 52,998 new sellers appearing in Month 2. As we use lagged reputation measures in our regressions, we are only able to use data for the 32,443 sellers who are in our data for at least two months.

is no less than 6 (or equivalently, her rating score is above 250; or a one-diamond status), and 0 otherwise. We choose this rating grade cutoff because one diamond is a well-known milestone status on Taobao. We then add the interaction of this dummy variable with the seller rating grade to the regressions specified by equation (1).

Table 8 presents the IV estimation results. The estimated coefficient of the rating grade indicates that a higher rating grade decreases the average revenue per transaction, increases the transaction volume, and has no significant effect on the monthly revenue. This set of results is consistent with our previous findings for different cohorts of new sellers. More importantly, the results in Table 8 show that the coefficients of the interaction term in all regressions are positive, implying that the effect of reputation on prices, transaction volume and monthly revenue changes when new sellers establish themselves: they cut prices less, sell even larger volumes and earn more monthly revenue as their ratings increase. In other words, although growing reputation still gives a new seller incentives to manage her reputation, she starts to enjoy some return to reputation as she becomes somewhat established. In Appendix B, we show the same pattern for a cohort of new sellers who first appear in our data in Month 3 (May 2010). We consider the results of these fixed-cohort analyses as strong support for the dynamic effects of reputation. These results also support the existence of reputation management as predicted by the theory on reputation dynamics.

Table 8: IV Results of the Fixed-Cohort Analysis

	log(Revenue)	log(1+Trans.)	log(Avg Revenue per Trans.)
L. Rating Grade	-0.072	0.556***	-0.266***
	(0.181)	(0.087)	(0.055)
L. Rating Grade$\times\mathbb{1}$(L. Rating Grade$\geq$6)	0.133***	0.063***	0.031***
	(0.029)	(0.014)	(0.008)
L. Rating Score	-1.167e-3	-3.352e-3***	1.239e-3***
	(9.403e-4)	(4.528e-4)	(2.641e-4)
L. % Positive Ratings	0.015	-0.025***	0.020***
	(0.011)	(0.005)	(0.003)
Seller fixed effects	yes	yes	yes
Month fixed effects	yes	yes	yes
Obs. (Seller/months)	222,935	222,935	198,623
Sellers	32,443	32,443	32,443

Note: This table reports the instrumental variable estimation results. ***: significant at the 1 percent level.

4.3 Survival Analysis

We have established that seller reputation has very different effects on sellers' short-term business performance for new sellers and established sellers. In this section, we investigate the role of a seller's reputation on a seller's survival likelihood. Specifically, we take a snapshot of the data and study how a seller's ratings at the time affect her survival likelihood six or twelve months later. As before, we estimate the effects of reputation for new sellers and established sellers separately.

Because we define new sellers as those who start selling in Month 2 or later, we take a Month-2 snapshot of our data so that we can study the effect of reputation for both groups of sellers. (In a robustness analysis in Appendix B, we use a Month-3 snapshot.) For the new sellers and established sellers in our data in Month 2, we examine their Month-8 and Month-14 survival outcomes. Among the 104,138 established sellers in our data in Month 2, 91.94% are still in our data in Month 8 and 79.16% in Month 14. By contrast, for the 52,998 new sellers in our data in Month 2, the survival rates are much lower. The 6-month survival rate and 12-month survival rate for these new sellers are 33.52% and 21.73%, respectively. This is not surprising because established sellers on average are more successful and have been selling for a longer period of time. Such a selection explains the different (level of) survival likelihood between these two groups of sellers. What about the slope of the survival probability with respect to reputation? Is the effect of reputation on survival also different between these two groups? To answer this question, we use a linear probability model to study how a seller's ratings affect her survival likelihood. Specifically,

$$\begin{aligned} Survival_i &= \beta_1(RatingGrade)_i + \beta_2(RatingScore)_i + \beta_3(\%PositiveRating)_i \\ &\quad + \gamma \boldsymbol{X}_i + \eta_{location\_trade} + \nu_i \end{aligned} \tag{2}$$

In equation (2), $Survival_i$ is a dummy variable equal to 1 if a seller i is still in the data six (twelve) months later, and 0 otherwise.[36] We use the same set of seller reputation variables as in equation (1), measured in Month 2. But we cannot incorporate seller fixed effects in this cross-sectional regression. Instead, we include seller location-trade fixed effects $\eta_{location\_trade}$, where location is a seller's residing city, and trade is her main business category. We also include the following observable seller attributes: age, gender, and whether the seller immigrated from her

[36] Admittedly, this is not a perfect measure of survival. For example, in the event that a seller takes a break in Month 14, the last month of our sample, but continues her business after, we would assign 0 to this seller's 12-month survival dummy. We have investigated this issue for our 6-month survival definition. We found that sellers who are not in the data 6 months later never appear in the data again.

birth province to her residence province. These seller attributes are denoted by $\boldsymbol{X}_i$ in equation (2). The error term, ν_i, is assumed to be *i.i.d.* across sellers.

Although we separate new sellers from established ones, sellers within each group may be heterogeneous along some important unobserved dimensions. We use the same instrumental variables as we have used for equation (1), but measured in Month 2, to deal with this endogeneity issue. Since we cannot incorporate seller fixed effects in this cross-sectional regression, the identification assumptions on the instrumental variables are stronger than those made for the main regressions. Therefore, one should read our results below with this caveat in mind.

Table 9 presents the IV estimation results. Our findings show that established sellers with higher ratings also have a higher survival likelihood. An increase in the rating grade by one grade increases a seller's 6-month survival likelihood by 5% and her 12-month survival likelihood by 9%. That is, a high reputation earns established sellers both short-run and long-run returns. By contrast, we find an insignificant negative effect of reputation on a new seller's survival likelihood. This result is consistent with our previous finding that new sellers manage their reputation actively by engaging in promotions at the cost of short-run returns. New sellers with higher ratings are more aggressive at doing so. Consequently, even though this "losing to win" strategy may push a few sellers to the top of the reputation ladder, it decreases sellers' chance of survival on average.

Why would sellers manage reputation so aggressively even at the cost of their survival likelihood? When the reputation premium that a seller can enjoy if she survives outweighs the adverse effect of a decreased survival likelihood, it is indeed rational for a seller – even a seller who is not risk-loving – to do so. Note that a new shop owner may have comparable products, services, and perhaps better prices than an established seller does, but she has little reputation to catch buyers' attention and to earn their trust. As reputation can only be accumulated via transactions, a new seller has an incentive to boost sales and earn reputation even at the cost of a short-run loss in monthly revenue or a long-run decline in business survival likelihood. These findings echo those of Foster, Haltiwanger and Syverson (forthcoming), who show that new plants are just as technically efficient as older plants, but start with considerably lower demand. In their paper, there is a "demand accumulation" process, such as building a customer base; in this paper, we suggest that there is a "reputation accumulation" process which new sellers go through to catch up with established sellers.

Table 9: The Effect of Seller Reputation on Survival

	Established Sellers		New Sellers	
	6 months	12 months	6 months	12 months
Rating Grade	0.047***	0.092***	-0.310	-0.233
	(0.006)	(0.009)	(0.631)	(0.530)
Rating Score	-1.584e-6	-4.280e-6**	9.871e-3	3.283e-3
	(1.249e-6)	(1.857e-6)	(0.059)	(0.050)
% Positive Ratings	0.011***	0.023***	0.003***	0.003***
	(0.002)	(0.003)	(0.001)	(0.001)
User Age	1.650e-4*	-2.358e-4	0.005***	0.005***
	(9.970e-5)	(1.490e-4)	(0.001)	(0.001)
User Gender	0.001	0.005**	-0.005	0.001
	(0.002)	(0.003)	(0.007)	(0.005)
Immigrant	-1.929e-4	0.004	0.027	0.030*
	(0.002)	(0.003)	(0.021)	(0.018)
Location/trade fixed effect	yes	yes	yes	yes
Obs. (Sellers)	104,138	104,138	52,998	52,998

Note: This table reports the instrumental variable estimation results for the regression specified in equation (2). ***: significant at the 1 percent level; **: significant at the 5 percent level; *: significant at the 10 percent level.

5 Conclusion

As Cabral and Hortacsu (2010) note, the "...eBay reputation system gives way to noticeable strategic responses from both buyers and sellers." Indeed, any reputation system may elicit strategic responses from both sides of the market. However, much of the previous empirical literature on reputation in both online and offline markets focuses on the behavior of buyers instead of that of sellers. Thus, our work contributes to the literature by studying the strategic responses from the seller side of a large-scale online retail market. Moreover, we study the dynamic effects of reputation. Guided by our simple model of reputation dynamics, we distinguish new sellers from established sellers and find that reputation has distinctively different effects on these two groups of sellers. Specifically, using a large panel of online sellers on China's leading e-commerce platform, Taobao.com, we find that established sellers receive substantial returns to reputation. Furthermore, consistent with the theory, we find evidence of reputation management for new sellersSuch reputation management behavior leads to a negative reputation premium in the short run and even decreases new sellers' survival likelihood in the longer run.

Additionally, this paper contributes to a better understanding of online entrepreneurship as we study the strategies and behavior of small, entrepreneurial online sellers. Most existing research focuses on established, mature firms. The evidence on how entrepreneurship originates and develops is rare as data on new business establishments are somewhat limited.[37] The detailed records we have obtained from Taobao provide valuable information on this topic. In particular, our study informs the design of online reputation systems as well as relevant policies in a new, thriving market environment. For example, our results suggest that platform design and credit environments to help new sellers survive the initial stages of reputation accumulation might play an important role in overcoming market inefficiencies due to informational frictions.

References

Ackerberg, Daniel, Gregory Crawford, and Jinyong Hahn (2011), "Orthogonal instruments: estimating price elasticities in the presence of endogenous product characteristics." Working paper, University of Warwick.

Akerlof, George A. (1970), "The market for "lemons": Quality uncertainty and the market mechanism." *Quarterly Journal of Economics*, 84, 488–500.

Anderson, Michael and Jeremy Magruder (2012), "Learning from the crowd: Regression discontinuity estimates of the effects of an online review database." *The Economic Journal*, 122, 957–989.

Bajari, Patrick and Ali Hortacsu (2004), "Economic insight from internet auctions." *Journal of Economic Literature*, 42, 457–486.

Bar-Isaac, Heski and Steven Tadelis (2008), "Seller reputation." *Foundations and Trends in Microeconomics*, 4, 273–351.

Boston Consulting Group (2012), *The Internet Economy in the G-20.*

Brown, Jennifer and John Morgan (2006), "Reputation in online markets: Some negative feedback." *California Management Review*, 49, 61–81.

Cabral, Luis and Ali Hortacsu (2010), "The dynamics of seller reputation: Evidence from eBay." *Journal of Industrial Economics*, 58, 54–78.

[37]Exceptions are studies that use data from the Census Bureau, for example, Haltiwanger, Jarmin and Miranda (2013).

Cai, Hongbin, Ginger Zhe Jin, Chong Liu, and Li-an Zhou (2014), "Seller reputation: From word-of-mouth to centralized feedback." *International Journal of Industrial Organization*, 34, 51 – 65.

Chevalier, Judith and Glenn Ellison (1999), "Career concerns of mutual fund managers." *Quarterly Journal of Economics*, 114, 389–432.

Chu, Junhong and Puneet Manchanda (forthcoming), "quantifying cross and direct network effects in online c2c platforms." *Marketing Science*.

Dellarocas, Chrysanthos (2003), "The digitization of word of mouth: Promise and challenges of online feedback mechanisms." *Management Science*, 49, 1407–1424.

Dobbs, Richard, Yougang Chen, Gordon Orr, James Manyika, Michael Chui, and Elsie Chang (2013), "China's e-tail revolution: Online shopping as a catalyst for growth." Mckinsey global institute report, McKinsey & Company.

Economist (2006), "China's pied piper." *Economist*, September 21.

Economist (2013), "The world's greatest bazaar; alibaba." *Economist*, March 23.

Einav, Liran, Theresa Kuchler, Jonathan Levin, and Neel Sundaresan (2015), "Assessing Sale Strategies in Online Markets Using Matched Listings." *American Economic Journal: Microeconomics*, 7, 215–47.

Elfenbein, Dan, Ray Fisman, and Brian McManus (2012), "Charity as a substitute for reputation: Evidence from an online marketplace." *Review of Economic Studies*, 79, 1441–1468.

Foster, Lucia, John C. Haltiwanger, and Chad Syverson (forthcoming), "The slow growth of new plants: Learning about demand?" *Economica*.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda (2013), "Who Creates Jobs? Small versus Large versus Young." *The Review of Economics and Statistics*, 95, 347–361.

Holmstrom, Bengt (1999), "Managerial incentive problems: A dynamic perspective." *Review of Economic Studies*, 66, 169–182.

Hong, Harrison, Jeffrey D. Kubik, and Amit Solomon (2000), "Security Analysts' Career Concerns and Herding of Earnings Forecasts." *RAND Journal of Economics*, 31, 121–144.

Houser, Dan and John Wooders (2006), "Reputation in auctions: Theory and evidence from eBay." *Journal of Economics and Management Strategy*, 15, 353–369.

Jolivet, Gregory, Bruno Jullien, and Fabien Postel-Vinary (forthcoming), "Reputation and prices on the e-market: Evidence from a major french platform." *International Journal of Industrial Organization*.

Ju, Heng, Liangjun Su, and Pai Xu (2013), "Pricing for goodwill: A threshold quantile regression approach." working paper, Singapore Management University.

Klein, Tobias J., Christian Lambertz, and Konrad O. Stahl (2015), "Market Transparency, Adverse Selection, and Moral Hazard." *Journal of Political Economy*, 1–49.

Luca, Michael (2011), "Reviews, reputation, and revenue: The case of Yelp.com." working paper, Harvard University.

Mayzlin, Dina, Yaniv Dover, and Judith Chevalier (2014), "Promotional reviews: An empirical investigation of online review manipulation." *American Economic Review*, 104, 2421–2455.

Olley, G Steven and Ariel Pakes (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64, 1263–97.

Resnick, Paul and Richard Zeckhauser (2002), "Trust among strangers in internet transactions: Empirical analysis of eBays reputation system." In *The Economics of the Internet and E-commerce* (Michael R. Baye, ed.), volume 11 of Advances in Applied Microeconomics, 127–157, Elsevier Science, Amsterdam.

Resnick, Paul, Richard Zeckhauser, John Swanson, and Kate Lockwood (2006), "The value of reputation on ebay: A controlled experiment." *Experimental Economics*, 9, 79–101.

Saeedi, Maryam (2014), "Reputation and adverse selection: Theory and evidence from eBay." working paper, Ohio State University.

Shapiro, Carl (1983), "Premiums for high quality products as returns to reputations." *Quarterly Journal of Economics*, 98, 659–680.

Ye, Qiang, Min Xu, Melody Kiang, Weifang Wu, and Fangfang Sun (2013), "In-depth analysis of the seller reputation and price premium relationship: A comparison between ebay US and Taobao China." *Journal of Electronic Commerce Research*, 14, 1–10.

Appendices

A First-Stage Results

Table A.1 presents the first-stage results of our analyses.

Table A.1: First-Stage Results

	New Sellers		Established Sellers	
	L. Rating Score	L. Rating Grade	L. Rating Score	L. Rating Grade
L. Buyer Transaction Grade	53.649***	0.559***	104.049***	0.178***
	(0.603)	(0.002)	(18.288)	(0.001)
L. Buyer Transactions	0.081***	6.165e-5***	3.813***	6.703e-5***
	(0.002)	(6.357e-6)	(0.080)	(4.588e-6)
L. % Positive Ratings	1.514	49.583***	13.347	48.566***
	(5.060)	(14.080)	(30.706)	(17.708)
Seller fixed effects	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes
F statistic	7939.27	81691.36	2016.36	43002.80
Obs. (Seller/months)	1,349,124		1,234,176	
Sellers	329,169		104,138	

***: significant at the 1 percent level.

B Robustness Analyses

B.1 Different Definitions of New Sellers and Established Sellers

In this section, we present the results of our analyses using different definitions of new sellers and established sellers. For our baseline results, we define new sellers and established sellers according to when a seller first appears in our data (referred to as the "first month criterion") and the cumulative seller transactions when she first appears in the data (referred to as the "transaction volume criterion"). Specifically, a new seller is a seller whose first appearance is later than Month 1 (March 2010) and who has no more than 30 transactions by then according to our baseline definitions. In Table B.2, we repeat the regression of log(Revenue) using the following alternative definitions of new sellers.

- Baseline Definition: first month > 1, cumulative transactions in the first month ≤ 30
- Definition 1: first month > 1, cumulative transactions in the first month ≤ 10
- Definition 2: first month > 1, cumulative transactions in the first month ≤ 40

- Definition 3: first month > 2, cumulative transactions in the first month ≤ 30
- Definition 4: cumulative transactions in the first month ≤ 30

In Definitions 1 and 2, we modify the transaction volume criterion (but keep the first month criterion). Specifically, in Definition 1 and Definition 2, the transaction volume criterion is 10 (the upper bound for grade 1) and 40 (the upper bound for grade 2), respectively, rather than 30 (as in the baseline definition). For Definition 3, we use a different first month criterion: a new seller's first month is later than Month 2. Finally, for Definition 4, we drop the first month criterion and use only the transaction volume criterion. The results in Table B.2 show that our findings are robust to these different definitions of new sellers.

Table B.2: Robustness Analyses: Different Definitions for New Sellers

Dep. Variable log(Revenue)	(1) Baseline	(2) Definition 1	(3) Definition 2	(4) Definition 3	(5) Definition 4
L. Rating Grade	-0.166***	-0.219***	-0.152***	-0.174***	-0.115***
	(0.038)	(0.046)	(0.039)	(0.040)	(0.039)
L. Rating Score	-1.152e-3***	-4.015e-4	-1.303e-3***	-1.321e-3***	-8.718e-4***
	(3.386e-4)	(4.795e-4)	(3.445e-4)	(3.774e-4)	(3.282e-4)
L. % Positive Ratings	0.019***	0.019***	0.018***	0.019***	0.020***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Seller fixed effects	yes	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes	yes
Obs. (Seller/months)	1,349,124	910,775	1,431,555	1,126,189	1,780,347
Sellers	329,169	231,998	346,590	296,726	382,623

***: significant at the 1 percent level.

Similarly, we report the robustness results using different definitions for established sellers in Table B.3. The definitions used are

- Baseline Definition: first month=1, cumulative transactions in the first month$\geq$251
- Definition 1: first month = 1, cumulative transactions in the first month$\geq$10001
- Definition 2: cumulative transactions in the first month$\geq$251

In Definition 1 (Column (2)), we modify the transaction volume criterion to be 10,001, the lower bound for the crown status (rather than 251, as in the baseline definition, which is the lower bound for the diamond status.) In Definition 2 (Column (3)), we again drop the first month criterion in defining an established seller and use only the transaction volume criterion. Again, the estimation results are robust.

Table B.3: Robustness Analyses: Different Definitions for Established Sellers

Dep. Variable log(Revenue)	(1) Baseline	(2) Definition 1	(3) Definition 2
L. Rating Grade	0.378***	1.169***	0.237***
	(0.039)	(0.288)	(0.039)
L. Rating Score	4.230e-5***	8.424e-7	-7.004e-5***
	(7.640e-6)	(2.938e-6)	(1.195e-5)
L. % Positive Ratings	0.337***	0.461***	0.194***
	(0.016)	(0.054)	(0.017)
Seller fixed effects	yes	yes	yes
Month fixed effects	yes	yes	yes
Obs. (Seller/months)	1,234,176	72,627	1,359,501
Sellers	104,138	5,678	130,311

***: significant at the 1 percent level.

B.2 Different Cohort for the Fixed Cohort Analysis in Section 4.2

In the fixed-cohort analysis in Section 4.2, we focus on new sellers who first appear in our data in Month 2 (April 2010) and find that the effect of ratings is different as these sellers become somewhat established. In this section, we run the same regression for new sellers who first appear in our data in Month 3 and report the estimation results in Table B.4. Comparing the findings in Table 8 and those in Table B.4, we see that our findings are robust to cohort choice. Specifically, the signs of all estimates are the same, except that for the effect of the lagged rating score on log(revenue), which is statistically insignificant in both regressions.

C An Illustrative Theoretical Model

In this section, we use a simple model to show that the price that a seller charges and the revenue that she earns can decrease with her rating in the early stage of her business, but increase with her rating later. The model consists of two periods in which a seller chooses prices to maximize total revenue in the two periods.[38] In the model, the seller's price decision in the first period has a dynamic implication for her revenue in the second stage.

Let $Q(p, r)$ be the demand in either period, which is decreasing in price p and increasing in rating r. At the beginning of the first period, a seller is endowed with a rating r_1. Her rating at the beginning of the second period depends on her rating in the first period (r_1) and her sales in

[38] We assume, for simplicity, that marginal cost is zero so that profit equals revenue and that there is no discounting.

Table B.4: Robustness Analysis: Different Cohort for the Fixed-Cohort Analysis (New Sellers who First Appear in Month 3)

	log(Revenue)	log(1+Trans.)	log(Avg Revenue per Trans.)
L. Rating Grade	0.060	0.187***	-0.069**
	(0.097)	(0.035)	(0.028)
L. Rating Grade$\times\mathbb{1}$(L. Rating Grade$\geq$6)	0.107***	0.025**	0.035***
	(0.030)	(0.011)	(0.008)
L. Rating Score	-1.712e-3***	-1.622e-3***	4.635e-4***
	(5.379e-4)	(1.934e-4)	(1.481e-4)
L. % Positive Ratings	0.020***	0.002	0.005***
	(0.006)	(0.002)	(0.002)
Seller fixed effects	yes	yes	yes
Month fixed effects	yes	yes	yes
Obs. (Seller/months)	181,260	181,260	164,074
Sellers	29,579	29,579	29,579

***: significant at the 1 percent level. **: significant at the 5 percent level.

the first period (q_1). Specifically, $r_2 = r_1 + \lambda q_1$, where λ is the average contribution of a seller's transaction to her rating. We assume that $\lambda > 0$.[39]

Let $\pi(p_t, r_t) = p_t Q(p_t, r_t)$ for $t = 1, 2$ be the seller's revenue in each period for a given price (p_t) and a given rating (r_t) in the corresponding period. In the second period, the seller chooses p_2 to maximize $\pi(p_2, r_2)$. Let $p_2^*(r_2)$ be the optimal price in the second period, and $\pi_2^*(r_2) = \pi(p_2^*(r_2), r_2)$ be the corresponding optimal revenue in the second period. In the first period, the seller chooses p_1 to maximize the sum of revenues from both periods, anticipating its effect on her rating next period. In other words, her problem in the first period is

$$\begin{aligned} \max_{p_1} V(p_1, r_1) &= \max_{p_1} \pi(p_1, r_1) + \pi_2^*(r_1 + \lambda Q(p_1, r_1)) \\ &= \max_{p_1} p_1 Q(p_1, r_1) + \pi_2^*(r_1 + \lambda Q(p_1, r_1)). \end{aligned} \tag{C.1}$$

C.1 The Effect of r_2 on the Second-period Price

If $\frac{\partial^2 \pi(p_2^*, r_2)}{\partial p_2 \partial r_2} = \frac{\partial Q_2}{\partial r_2} + p_2^* \frac{\partial^2 Q_2}{\partial p_2 \partial r_2} > 0$, a standard comparative statics argument implies that the optimal second-period price is increasing in r_2. Since $\frac{\partial Q_2}{\partial r_2} > 0$, a sufficient condition is that $\frac{\partial^2 Q_2}{\partial p_2 \partial r_2} \geq 0$, i.e., a seller's demand becomes less sensitive to a change in price, or remains the same,

[39] Our results below rely on this assumption. As explained in the introduction, on Taobao, each transaction contributes to a seller's rating by -1, 0, or +1. Therefore, λ can be negative in principle. But in the data, we find that conditional on a positive transaction volume last month, 99.56% of seller-month observations have a positive growth in the seller rating.

as r increases. Under this condition, $\frac{\partial^2 \pi(p_2^*, r_2)}{\partial p_2 \partial r_2} > 0$, and hence, the optimal p_2 is increasing in r_2. This is intuitive. Since an increase in ratings shifts the demand and decreases consumers' sensitivity to price, it is optimal for the seller to charge a higher price.

C.2 The Effect of r_2 on Second-period Revenue

By the Envelope Theorem, we have

$$\frac{d\pi_2^*}{dr_2} = p_2^* \frac{\partial Q_2}{\partial r_2} > 0, \tag{C.2}$$

i.e., the seller's second-period revenue increases with her rating in the second period.

Moreover,

$$\begin{aligned} \frac{d^2 \pi_2^*}{dr_2^2} &= \frac{\partial^2 \pi\left(p_2^*, r_2\right)}{\partial r_2^2} + \frac{\partial^2 \pi\left(p_2^*, r_2\right)}{\partial p_2 \partial r_2} \frac{dp_2^*}{dr_2} \\ &= p_2^* \frac{\partial^2 Q_2}{\partial r_2^2} - \left(\frac{\partial^2 \pi\left(p_2^*, r_2\right)}{\partial p_2 \partial r_2} \right)^2 / \frac{\partial^2 \pi\left(p_2^*, r_2\right)}{\partial p_2^2}, \end{aligned} \tag{C.3}$$

where the second equality holds because $\frac{dp_2^*}{dr_2} = -\frac{\partial^2 \pi(p_2^*, r_2)}{\partial p_2 \partial r_2} / \frac{\partial^2 \pi(p_2^*, r_2)}{\partial p_2^2}$. The second term in (C.3) is positive by the second-order condition. The first term in (C.3) contains $\frac{\partial^2 Q_2}{\partial r_2^2}$. When the marginal effect of ratings does not diminish too fast as ratings increase, i.e., when either $\frac{\partial^2 Q_2}{\partial r_2^2} \geq 0$ or $\frac{\partial^2 Q_2}{\partial r_2^2} < 0$ but its absolute value is not too large, $\pi_2^*\left(r_2\right)$ is not only increasing in r_2 but also convex in r_2. For example, if demand $Q\left(p, r\right)$ is linear in r, the first term in (C.3) is zero and therefore $\frac{d^2 \pi_2^*}{dr_2^2} > 0$. As we will show in the next subsection, the convexity of $\pi_2^*\left(r_2\right)$ has a significant implication for the effect of r_1 on the optimal p_1 in the first period.

C.3 The Effect of r_1 on the First-period Price

The first-order condition for the seller's first-period problem (C.1) is

$$\frac{\partial V}{\partial p_1} = p_1 \frac{\partial Q_1}{\partial p_1} + Q_1 + \lambda \frac{d\pi_2^*}{dr_2} \frac{\partial Q_1}{\partial p_1}, \tag{C.4}$$

From the first-order condition (C.4), we can see that in addition to the standard static tradeoff between the profit margin and demand in the current period, the seller also considers how the current price affects future revenue through affecting the ratings next period. In particular, because $\frac{d\pi_2^*}{dr_2} > 0$ as shown in (C.2) and $\frac{\partial Q_1}{\partial p_1} < 0$ under our assumption on the demand function, this dynamic effect is negative. In other words, when the seller lowers the price in the first period, she sells more in the first period, which helps her accumulate ratings and thus affects her demand and revenue in the second period.

To see how the seller's optimal first-period price varies with her rating in the first period, note that

$$\begin{aligned}\frac{\partial^2 V}{\partial p_1 \partial r_1} &= \frac{\partial \pi\left(p_1, r_1\right)}{\partial p_1 \partial r_1}+\frac{\partial \pi_2^*\left(r_1+\lambda Q\left(p_1, r_1\right)\right)}{\partial p_1 \partial r_1} \\ &= p_1 \frac{\partial^2 Q_1}{\partial p_1 \partial r_1}+\frac{\partial Q_1}{\partial r_1}+\frac{d^2 \pi_2^*}{d r_2^2}\left(1+\lambda \frac{\partial Q_1}{\partial r_1}\right) \lambda \frac{\partial Q_1}{\partial p_1}+\frac{d \pi_2^*}{d r_2} \lambda \frac{\partial^2 Q_1}{\partial p_1 \partial r_1}.\end{aligned} \tag{C.5}$$

In equation (C.5), the signs of the first and the last terms depend on $\frac{\partial^2 Q_1}{\partial p_1 \partial r_1}$, the sign of which is ambiguous. The second term is positive. In the third term, $\left(1+\lambda \frac{\partial Q_1}{\partial r_1}\right) \lambda \frac{\partial Q_1}{\partial p_1}<0$. As we have shown in Section C.2, under certain conditions, $\frac{d^2 \pi_2^*}{d r_2^2}>0$, in which case the third term above is negative. When $\frac{d^2 \pi_2^*}{d r_2^2}$ is sufficiently large, the negative third term may lead to a negative $\frac{\partial^2 V}{\partial p_1 \partial r_1}$, implying that $p_1^*\left(r_1\right)$ is decreasing in r_1.

To understand the intuition, let us look at a simple linear example. Suppose $Q(p, r)=-\theta_p p+\theta_r r$ where $\theta_p, \theta_r>0$. In this example, $p_2^*\left(r_2\right)=\frac{\theta_r}{2 \theta_p} r_2$, and $\pi_2^*\left(r_2\right)=\frac{\theta_r^2}{4 \theta_p} r_2^2$, which is convex in r_2. The cross derivative in (C.5) becomes

$$\frac{\partial^2 V}{\partial p_1 \partial r_1}=\theta_r-\frac{\theta_r^2}{2 \theta_p}\left(1+\lambda \theta_r\right) \lambda \theta_p=\theta_r-\frac{1}{2}\left(1+\lambda \theta_r\right) \lambda \theta_r^2 \tag{C.6}$$

The first term $\left(\theta_r\right)$, which is $\frac{\partial \pi\left(p_1, r_1\right)}{\partial p_1 \partial r_1}$, captures how marginal revenue in the first period (from increasing p_1) changes with r_1. For a seller with a higher rating r_1, this marginal revenue from increasing the price is larger as a higher rating implies a larger demand for any given price. The second term, $-\frac{1}{2}\left(1+\lambda \theta_r\right) \lambda \theta_r^2$, which is $\frac{\partial \pi_2^*\left(r_1+\lambda Q\left(p_1, r_1\right)\right)}{\partial p_1 \partial r_1}$, is negative and captures how marginal revenue in the second period (from decreasing p_1) changes with r_1. Since $\pi_2^*\left(r_2\right)$ is convex in this example, for a seller with a higher rating r_1 (and hence a higher rating r_2 *ceteris paribus*), the marginal effect of increasing p_1 on π_2^* through affecting r_2 is larger. Our result indicates that when $\lambda \theta_r>1$, the second term (negative) dominates the first term (positive). Consequently, $\frac{\partial^2 V}{\partial p_1 \partial r_1}<0$ and hence, $p_1^*\left(r_1\right)$ is decreasing in r_1. Intuitively, when ratings have a large impact on demand (i.e., a large θ_r) and when current sales have a large impact on the future rating (i.e., a large λ), the dynamic effect captured by the second term dominates the first static effect. As a result, $p_1^*\left(r_1\right)$ is decreasing in r_1.

C.4 The Effect of r_1 on First-period Revenue

To see how r_1 affects the first-period revenue, we compute the derivative of the first-period revenue with respect to r_1 at the optimal price $p_1^*(r_1)$. Let $\pi_1^*(r_1) = p_1^*(r_1) Q(p_1^*(r_1), r_1)$. Then,

$$\frac{d\pi_1^*(r_1)}{dr_1} = p_1^* \frac{\partial Q_1}{\partial r_1} + \left(p_1^* \frac{\partial Q_1}{\partial p_1} + Q_1\right) \frac{dp_1^*}{dr_1}, \tag{C.7}$$

where the first term captures the direct positive effect of r_1 on the first-period revenue, and the second term captures the indirect effect of r_1 on the revenue through affecting p_1^*. The first-order condition (C.4) implies $\left(p_1^* \frac{\partial Q_1}{\partial p_1} + Q_1\right) = -\frac{d\pi_2^*}{dr_2} \lambda \frac{\partial Q_1}{\partial p_1} < 0$. Therefore, when $\frac{dp_1^*}{dr_1} < 0$, the second term in (C.7) is negative. When the second term dominates the first term, π_1^* is decreasing in r_1.

In our linear example, $p_1^*(r_1) = \frac{[2-\lambda\theta_r(1+\lambda\theta_r)]\theta_r}{(4-\lambda^2\theta_r^2)\theta_p} r_1$.[40] Note that the second-order condition implies that $\lambda\theta_r < 2$,[41] so that $p_1^* > 0$. We can write (C.7) in this linear example as

$$\begin{aligned} \frac{d\pi_1^*(r_1)}{dr_1} &= [\theta_r r_1 + (-\theta_p p_1^* - \theta_p p_1^* + \theta_r r_1)] \frac{dp_1^*}{dr_1} \\ &= 2\theta_r r_1 \left(1 - \frac{2 - \lambda\theta_r(1+\lambda\theta_r)}{4 - \lambda^2\theta_r^2}\right) \frac{dp_1^*}{dr_1} \\ &= 2\theta_r r_1 \left(\frac{2 + \lambda\theta_r}{4 - \lambda^2\theta_r^2}\right) \frac{dp_1^*}{dr_1}. \end{aligned} \tag{C.8}$$

Therefore, when $\frac{dp_1^*}{dr_1} < 0$, the revenue in the first period is also decreasing in r_1 in this example.

In summary, our simple model presents a situation in which a seller faces a dynamic tradeoff in setting prices. Lower prices in the first period lead to higher ratings in the second period, and thus higher payoffs in the second period. Moreover, such a dynamic effect of lowering prices in the first period can be increasing with the seller's rating in the first period. This is because both a decrease in price and an increase in ratings in the first period can positively contribute to an increase in the period-2 rating. When the optimal period-2 revenue is convex in period-2 rating, the effect of the lowering period-1 price increases with the period-1 rating. As a result, the seller's optimal price (as well as revenue) in the first period decreases with her first-period rating, while her optimal price (as well as revenue) in the second period increases with her second-period rating. This simple model shows that a seller's reputation management behavior in the early stages of her business may lead

[40] To see this, note that the first-order condition (C.4) in our linear example is $-\theta_p p_1 - \theta_p p_1 + \theta_r r_1 + \frac{\theta_r^2}{2\theta_p}(r_1 + \lambda(-\theta_p p_1 + \theta_r r_1))\lambda(-\theta_p) = 0$.

[41] To see this, note that $V(p_1, r_1) = p_1(-\theta_p p_1 + \theta_r r_1) + \frac{\theta_r^2}{4\theta_p}(r_1 + \lambda(-\theta_p p_1 + \theta_r r_1))^2$. Therefore, $\frac{\partial^2 V}{\partial p_1^2} = -\frac{\theta_p}{2}\left[4 - (\lambda\theta_r)^2\right] < 0$ if and only if $\lambda\theta_r < 2$.

to different effects of reputation over the life-cycle of her business, which is consistent with our empirical findings in Section 4.

C.5 Discussion of a More General Demand Function

In this subsection, we consider a more general demand function $Q(p, r, R)$, where p is price, r is the rating score and R is the rating grade, which is an increasing step function of r. In other words, for a given rating grade, demand is increasing in the rating score. When there is an increase in the rating grade, there is an additional increase in the demand. For example, to extend the linear example above, the generalized linear demand function is $Q(p, r, R) = -\theta_p p + \theta_r r + \theta_R R$. The transition of the rating score remains the same: $r_2 = r_1 + \lambda q_1$. The transition of the rating grade can be written as $R_2 = R_1 + G(\lambda q_1)$, where G equals 1 if λq_1 is above a certain cutoff and 0 otherwise.

In Section C.3, we show that conditional on the rating grade R_1, the marginal effect of the rating score r_1 on the optimal p_1^* depends on θ_r. When θ_r is relatively small, p_1^* is increasing in r_1; and when θ_r is relatively large, p_1^* is decreasing in r_1. Intuitively, when the rating score has a large impact on demand, the dynamic effect dominates the static effect, and vice versa. Even though R_1 is an integer,[42] the intuition developed in the above comparative statics also applies to the effect of R_1 on p_1^*. This is because R_1 and r_1 affect demand in a similar way (i.e., $Q(p, r, R) = -\theta_p p + \theta_r r + \theta_R R$) and their transitions are also similar (i.e., $r_2 = r_1 + \lambda q_1$ and $R_2 = R_1 + G(\lambda q_1)$, where G is an increasing function). Therefore, based on similar intuitions for the comparative statics on r_1, we expect the effect of R_1 on p_1^* to depend on how large θ_R is.

As we argue in the paper, the rating grade R represented by a system of hearts, diamonds, crowns and golden crowns is more salient than the rating score r. We expect θ_R to be larger than θ_r. Therefore, it is possible that while within a rating grade, a seller charges a higher price as her rating *score* increases (i.e., the static effect dominates the dynamic effect because θ_r is small), the seller charges a lower price as her rating *grade* increases (i.e., the dynamic effect dominates because θ_R is large), which is consistent with our finding in Section 4.1.

[42] Moreover, a change in R_1 is necessarily companied by a change in r_1 by at least one point. As we will argue later, we expect θ_r to be small. Therefore, in a comparative statics analysis of a change in R_1, it is not unreasonable to assume that the effect of the accompanying change in r_1 is dominated by the direct effect of the change in R_1.