Certification, Reputation and Entry: An Empirical Analysis *

Xiang Hui MIT

Maryam Saeedi CMU

Giancarlo Spagnolo SITE, Tor Vergata, EIEF & CEPR

Steve Tadelis† UC Berkeley, NBER & CEPR

February 5, 2018

Abstract

How does quality-certification affect product quality in markets? We exploit a policy change on eBay to analyze how a more stringent certification policy affects entry and behavior across many markets segments. We find that first, entry increased in markets where it was harder to get certified, until a new steady state was reached. Second, the quality distribution of entrants exhibits fatter tails after the policy, and overall quality is slightly higher. Last, some incumbents respond by increasing the quality of their service to maintain certification. The results inform the design of certification policies in electronic and other markets. *JEL* D47, D82, L15, L86

*We would like to thank Sven Feldmann, Hugo Hopenhayn, Brian McManus, Peter Newberry, Konrad Stahl, and many seminar and conference participants for excellent suggestions, discussions and comments.

†xianghui@mit.edu, msaeedi@andrew.cmu.edu, giancarlo.spagnolo@hhs.se, stadelis@berkeley.edu

1 Introduction

In his seminal paper, Akerlof [1970] showed that asymmetric information can cause adverse selection, inducing only low-quality sellers to enter a market. Market institutions have emerged to help mitigate adverse selection, including warranties (Grossman [1981]), reliance on past reputation (Shapiro [1983]) and regulated certification by a trusted institution. Online marketplaces employ all three in the form of buyer protection policies, seller reputation scores, and badges that certify sellers who meet some minimum quality threshold determined by the marketplace. Examples of such badges are eBay's "Top Rated Seller", Airbnb's "Superhost", and Upwork's "Rising Star".

Certification badges can alleviate some market frictions caused by asymmetric information, but at the same time they can become entry barriers for new high-quality entrants who do not have a certifiable track record (Klein and Leffler [1981], Grossman and Horn [1988]). Hence, different certification criteria will impact the perceived quality of sellers both with and without certification, and in turn, the market structure mix of incumbents and entrants. How will more stringent certification criteria impact the incentives of new sellers to enter the market? And how will it change the quality distribution of sellers in the market? Answering these questions sheds light on how the design of certification policies may affect the overall evolution of markets.

In this paper, we shed light on these questions by analyzing the effects of a change in the certification policy of eBay, one of the largest and best-known e-commerce markets. We exploit a quasi-experiment that occurred in 2009 when eBay replaced the "Powerseller" badge awarded to particularly virtuous sellers with the "eBay Top Rated Seller" (eTRS) badge that had more stringent requirements.

To guide our empirical analyses, we develop a simple theoretical model in which this policy change alters the incentives of potential entrants who differ in their quality. A more stringent badging requirement causes the average quality of both badged and unbadged sellers to increase, because the sellers who lose their badge are worse than those who remain badged, but are better than those who were not badged previously. Hence, the change may induce more entry of top-end-quality firms by increasing their future payoff from obtaining a more selective badge, but it may also induce more entry of low-quality entrants, because they will be pooled with better non-badged sellers. More average-quality sellers may find entry to be less attractive after the policy change if obtaining a badge becomes impossible, or may change their behavior and exert higher effort if they see that they can maintain the more selective badge. With time, the structure and quality of the seller population is affected, and new equilibria may be reached depending on market characteristics.

To identify these potential effects empirically, we exploit differential exposure to the policy change across 400 different subcategories of the eBay marketplace. This presents us with a quasi-experiment across different subcategories on eBay because the change in certification requirements is universal across all subcategories, but the difficulty to meet the new requirements exogenously differ by subcategories leading to heterogeneous effects of the policy in terms of fraction of badged sellers. We treat each subcategory as a separate market, and define the entry date of a seller in a particular market as the first date that the seller made a listing in that subcategory.

We first document a significant drop in the share of badged sellers at the policy change date, which is what the policy change was designed to do. We show that there is substantial heterogeneity of this effect across subcategories, consistent with the fact that the difficulty of obtaining the badge is exogenously different across markets. We then show that there is a negative correlation between the share of badged sellers and the number of entrants across subcategories, suggesting an entry deterrence effect of certification; and that after the policy change, this correlation becomes stronger. However, this change is temporary, it tends to disappear once the market adjusts to a "new equilibrium", which occurs after about six months.

Turning to our main analysis, we find that in the first three to six months after the policy change, entry increases more in markets that were affected more by the policy (where the fraction of badged sellers fell relatively more). A 10% larger drop in the fraction of badged sellers results in a 3% increase in entry. However, this effect becomes statistically insignificant when we consider a longer period of seven to twelve months after the policy change. We then show that the average quality provided by entrants increases significantly after the policy change. In contrast to the long-term effect of the policy change on the number of entrants, this effect on quality persists over a longer time period. We also find that the entrants in the more affected subcategories tend to be smaller on average; however, their total market share increases after the policy change.

Importantly, we find that the distribution of the quality provided by entrants also changes with the policy and exhibits fatter tails. In particular, a larger share of entrants provide quality at the top and bottom quintiles of the quality distribution. This finding is consistent with the prediction that sellers from the extremes of the quality distribution have stronger incentives to enter immediately after the policy change is implemented.

Aside from affecting the selection of entrants, an increase in quality could also be due to sellers changing their behavior and choosing to provide higher quality, suggesting that the change may help solve a moral hazard problem. We therefore study the behavior of different types of *incumbent* sellers–with and without a badge before and after the policy change–and find almost no change in their quality. Consistent with our model, the only group of incumbents that show a significant change in behavior linked to the policy are those that lose their badge and, improving quality provision, manage to re-gain it within three months. Taken together this suggests that moral hazard issues are present but are not entirely dominant, and that a significant part of the increase at the tails of the quality distribution we observe for entrants is linked to selection, besides changes in behavior.

We then study how price and market share changed for four groups of incumbent sellers, depending on their badge status before and after the change in policy. The results we find are intuitive: first, sellers who had badges and lost them experienced a decrease in the relative price that they receive. Second, sellers who were not previously badged but receive a badge after the change experienced the largest growth in market share.[1] Third, sellers who were or weren't badged before and after the policy change experience changes that are in between the other two groups.

Finally, we perform a series of robustness tests. First, we perform a placebo test that provides evidence consistent with the exclusion restriction in our econometric specification. Second, we perform our analyses for two types of entrants into a market: new sellers on eBay and existing sellers entering a new subcategory. The estimates across the groups are very similar. Third, we study how exits have changed and find that the quality distribution of exits has "thinner" tails, which is consistent with the policy change improving incumbents' outcomes at the tails. Last, we check the robustness of our results to several econometric specifications; the results are reported in the online appendix.

Our paper joins a growing literature that uses rich online marketplace data to understand how to foster trade and alleviate asymmetric information in markets. The closest papers to ours are Elfenbein et al. [2015], Klein et al. [2016], and Hui et al. [2017], which also used data from eBay to study the effects of different information policies on market structure. In particular, Elfenbein et al. [2015] studied the value of a certification badge across different markets among different types of sellers. They found that certification provides more value when the number of certified sellers is low and when markets are more competitive. They did not explicitly study the impact of certification

[1] The existence of sellers who were not badged before but are after the policy change is due to sellers not being badged instantaneously when they meet the certification requirements, but being instead certified once every month.

on the dynamics of entry and the changes in market structure.

Klein et al. [2016] and Hui et al. [2017] exploited a different policy change on eBay after which sellers could no longer leave negative feedback for buyers, reducing the costs for buyers of leaving negative feedback. Both studies found an improvement in buyers' experience after the policy change. Using scraped data, Klein et al. [2016] cleverly take advantage of the evolution of both the public feedback and the anonymous feedback of Detailed Seller Ratings (DSR) to show that the improvement in transaction quality is not due to exits from low-quality sellers. Using internal data from eBay, Hui et al. [2017] complement Klein et al. [2016] and further investigate changes in the size of incumbents. They found that although low-quality sellers do not exit after the policy change, their size shrinks dramatically, which accounts for at least 68% of the quality improvement. In comparison with these three papers, our paper explicitly studies the impact of certification on the dynamics of entry and the changes in market structure, as well as the quality provided by incumbents before and after the change.

Our paper also relates to the literature that analyzes the effects of changes in eBay's feedback mechanisms on price and quality (e.g. Klein et al. [2016], Hui et al. [2016], and Nosko and Tadelis [2015]). Consistent with the findings reported in these papers, we found that the sellers that are badged both before and after the policy change are of higher quality than sellers that were only badged before but not after the policy change. In addition, the sellers that are badged both before and after the policy change also benefit from higher conversion rates, because the new badge carries higher value than the old one. More generally, our paper also broadly relates to the literature that analyzes the effect of reputation and certification on sales performances, such as Chevalier and Mayzlin [2006], Chintagunta et al. [2010], Zhu and Zhang [2010], Zhao et al. [2013], Wu et al. [2015], Hui et al. [2016] and Proserpio and Zervas [2017]. (See Bajari and Hortacsu [2004], Dellarocas et al. [2006], Dranove and Jin [2010] and Tadelis [2016] for surveys.)

Our results have implications for the design of certification mechanisms in electronic markets, where a host of performance measures can be used to set certification requirements and increase buyers' trust in the marketplace. They may also offer useful insights for other markets with high levels of asymmetric information, such as in public procurement, where regulatory certification can significantly change the competitive environment and reduce the costs of public services.[2]

[2]For example, concerns have been expressed by several prominent U.S. senators and the EU that the extensive use of past performance information for selecting federal contractors could hinder the ability of new or small businesses to enter public procurement markets. The debate led the General Accountability Office to study dozens of procurement decisions across multiple government agencies, but the resulting report, published in 2011, was rather inconclusive (more discussions in Butler et al. [2013]).

The remainder of the paper is organized as follow. Section 2 provides details about the policy change. In Section 3 we provide a framework using a simple theoretical example to illustrate how the policy could affect entry. Section 4 describes our data, and Section 5 discusses our empirical strategy. In Section 6, we provide our results, while in Section 7, we provide robustness tests. Section 8 concludes the paper.

2 Background and Policy Change

eBay started with its well-studied feedback rating in which sellers and buyers can give one another a positive, negative, or neutral feedback rating. eBay then introduced "detailed seller ratings," in which buyers give sellers an anonymous rating between 1 and 5 stars in four subcategories (item as described; communication; shipping rate; and shipping speed). To combat concerns that retaliation prevents buyers from leaving honest negative feedback, in 2008 eBay made the feedback rating asymmetric so that sellers could only leave a positive or no rating for buyers.

In addition to user-generated feedback, eBay started certifying what it deemed to be the highest-quality sellers by awarding them the "Powerseller" badge. To qualify for the Powerseller program, a seller needed to sell at least 100 items or at least $1000 worth of items every month for three consecutive months. The seller also needed to maintain at least 98% of positive feedback and 4.6 out of 5.0 detailed seller ratings. Finally, a seller had to be registered with eBay for at least 90 days. The main benefit of being a Powerseller was receiving discounts on shipping fees of up to 35.6%. There were different levels of Powersellers depending on the number and value of annual sales, but all Powersellers enjoyed the same direct benefits from eBay. An indirect benefit of the Powerseller badge was that it made very salient that the badged seller had constantly been performing well and is therefore likely to be a higher quality seller.

eBay revised its certification requirements and introduced the "eBay Top Rated Seller" (eTRS) badge, which was announced in July 2009 and became effective in September 2009. To qualify as a Top Rated Seller, a seller must have the Powerseller status. Additionally, the seller needs to have at least 100 transactions and sell $3000 worth of items over the previous 12 months, and must have less than 0.5% or 2 transactions with low DSRs (1 or 2 stars), and low dispute rates from buyers (less than 0.5% or 2 complaints from buyers).[3] The information on dispute rates, only available to

[3] A Senior Director who was involved in the change explained that there were two main reasons for the change: First, the Powerseller program rewarded sellers with higher discounts on their final value fees based on their sales volume, mostly irregardless of their performance, which created an incentive for sellers to sell more, without considering the experience they were delivering. Second, buyers perceived the Powerseller badge to mean eBay endorsed the seller.

eBay, was not used before. It is also important to note that after the introduction of eTRS, sellers can still obtain the Powerseller status but it is no longer displayed as a badge for buyers to observe.

Top Rated Sellers must meet stricter requirements than previous Powersellers, but also enjoy greater benefits. Top Rated Sellers receive a 20% discount on their final value fee (a percent of the transaction price) and have their listings positioned higher on eBay's "Best Match" search results page, which is the default sorting order, and results in more sales. Finally, the Top Rated Seller badge appears on all listings from a Top Rated Seller, signaling the seller's superior quality to all potential buyers.

3 Certification and Entry: A Simple Framework

To guide our analysis, we present a simple model that incorporates both hidden information (adverse selection) and hidden action (moral hazard) using a three-type model in the spirit of Diamond [1989].

Supply: Consider a market with a continuum of sellers. Each seller can produce and sell one unit or output with zero marginal costs and fixed costs $k \in [0, \infty)$, which is independently distributed across all sellers with the continuous and strictly increasing cumulative distribution function $G(k)$, with $G(0) = 0$ and $G(\infty) = 1$. There are three types of sellers: a measure μ_ℓ of "low-quality" sellers, indexed by ℓ, who can only produce low quality L; a measure μ_h of "high-quality" sellers, denoted by h, who can only produce high-quality H; and a measure μ_s of strategic sellers, denoted s, who can each choose whether to exert effort at a cost e and produce high quality H, or whether to shirk at no cost and produce medium quality M, where $H > M > L \geq 0$. The cost of effort $e \in [0, \infty)$ is independently distributed across all s-type sellers with the continuous and strictly increasing cumulative distribution function $F(e)$, with $F(0) = 0$ and $F(\infty) = 1$. Hence, s-type sellers have two dimensions of cost heterogeneity, (k, e), while ℓ- and h-type sellers only differ across k.

Demand: We assume a continuum of buyers who each demand one unit of a good and are willing to pay up to the expected quality of the good. To simplify, we assume that the buyers are on the "long side" of the market so that market clearing prices leave buyers with no surplus and the price of each good will be equal to its expected quality.

This skewed purchasing towards Powersellers, who already had a pricing advantage over non-Powersellers due to the discounts, but had little incentive to deliver great service. The Top-Rated badge introduced minimum performance requirements to obtain discounts by using maximum thresholds of low DSRs and dispute rates.

Information: Buyers cannot observe the quality of any given seller. A marketplace regulator can, however, produce an observable "badge" $B \in \{M, H\}$ that credibly signals if a seller's quality is at least at the threshold B. Given a badge B, let $\underline{v}_B$ denote the expected quality of sellers who are below the badge threshold and let $\bar{v}_B$ denote the expected quality of sellers who are at or above the badge threshold. Note, therefore, that if a positive measure of sellers of all types are in the market, then $\bar{v}_H = H$ and $M > \underline{v}_H > L$, whereas $H > \bar{v}_M > M$ and $\underline{v}_M = L$.

Equilibrium: Let $\mu_{\theta B}$ denote the measure of type θ sellers that enter the market when the badge is B. Let π denote the fraction of active s-type sellers who choose to exert effort. An *equilibrium* for threshold $B \in \{M, L\}$ is a pair of prices $\underline{p}_B$ and $\bar{p}_B$, and measures of each type of sellers, $\mu_{\theta B}$, and the proportion of s-type sellers who enter and work π, such that prices equal expected qualities, which in turn are consistent with Bayes rule given the measures of sellers of each type above and below the threshold, and that all sellers are best responding to prices.

We are interested in the comparative statics of making the badge more restrictive by first having $B = M$, and then increasing the threshold to $B = H$ so that only the highest-quality sellers are awarded a badges.

3.1 Lax Badge: $B = M$

The following observation is straightforward:

Lemma 1. . *All s-types choose to shirk when $B = M$.*

The result is obvious: because $B = M$, all s-types qualify to be badged whether they choose to exert effort or not. Since prices depend only on the badge, there are no returns to effort while the cost of effort is positive for all s-types.

The equilibrium is now easy to characterize as follows:

1. **prices:** $\bar{p}_M = \bar{v}_M = \frac{\mu_s M + \mu_h H}{\mu_s + \mu_h}$, and $\underline{p}_M = \underline{v}_M = L$,
2. **entry:** $\mu_{\ell M} = G(L)\mu_\ell$, $\mu_{sM} = G(\bar{v}_M)\mu_s$ and $\mu_{hM} = G(\bar{v}_M)\mu_H$,
3. **behavior:** All s-types who enter choose to shirk.

The measure of sellers who enter are determined by those who can cover their fixed costs given the two equilibrium prices. $\bar{p}_M$ is equal to the expected quality given the weights of the s- and h-types in the population because $G(\cdot)$ is i.i.d. across all types, both $s-$ and h-types receive the same price, and have the same zero-profit condition.

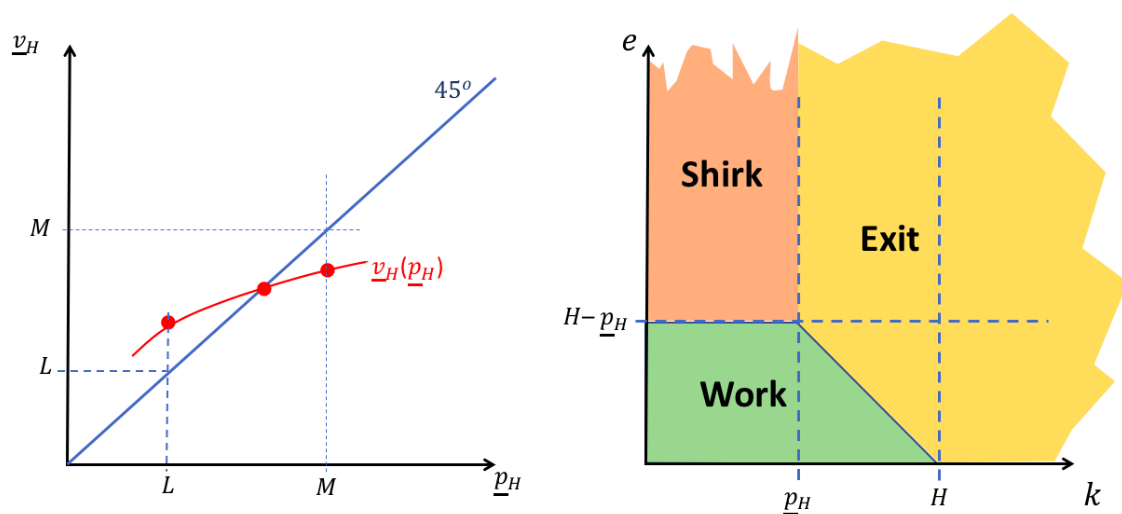
3.2 Stringent Badge: $B = H$

When $B = H$, s-type sellers will be badged if and only if they choose to exert effort.

Lemma 2. . $1 > \pi > 0$ *in any equilibrium with* $B = H$.

Proof. Because $G(\cdot)$ is continuous and increasing on $[0, \infty)$, and $G(0) = 0$, it follows that a positive measure of all types will enter the market. In any equilibrium with $B = H$, if a positive measure of all types enters, then $\bar{v}_H = H$ and $M > \underline{v}_H > L$. This implies that in equilibrium $\bar{p}_H > \underline{p}_H$. In turn, because both $F(\cdot)$ and $G(\cdot)$ are continuous and increasing on $[0, \infty)$, and $G(0) = F(0) = 0$, it follows that a positive measure of s-types will choose to prefer to enter, exert effort and be badged over not being badged. Finally, because $F(\infty) = 1$, and because $\bar{p}_H - \underline{p}_H$ is bounded, then not all s-types who enter will exert effort. $\square$

Figure 1: Equilibrium when $B = H$



To characterize equilibrium when $B = H$, it is illustrative to graphically describe the structure of any equilibrium as shown in Figure 1. The right panel shows the two-dimensional cost-space of s-type sellers who have both entry costs k and effort costs e. Because $\bar{p}_H = H$, it's clear that any s-type with $k > H$ cannot earn positive profits and will exit. Similarly, any s-type with $k < \underline{p}_H$ can enter and earn $\underline{p}_H - k > 0$ by shirking. For these entrants, the benefit from working outweighs the cost of working if and only if $H - \underline{p}_H > e$. Finally, for those with fixed costs $H > k > \underline{p}_H$, if $k + e < \underline{p}_H$ then they prefer to enter and work over exit (note that shirking yields negative profits), while if $k + e > \underline{p}_H$ then they prefer to enter. This observation helps characterize the equilibrium as follows:

Proposition 1. *When $B = H$ there exists an equilibrium with $\bar{p}_H = H$ and $M > \underline{p}_H > L$.*

Proof. Market prices determine which types enter and whether entering s-types choose to work or shirk, and by construction, $\bar{p}_H = H$. Consider the lowest possible unbadged price, $\underline{p}_H = L$. Because $L > 0$, a proportion $G(L)$ of ℓ- and s-types with fixed costs $k < L$ will enter, of which a proportion $\pi = F(H - L)$ of s-types will work and obtain a badge, and the remainder will shirk and produce quality M. But because a positive measure $G(L)(1 - F(H - L))$ of s-types enter and are unbadged, it follows that $\underline{v}_H > \underline{p}_H = L$, so this cannot be an equilibrium. Now define $\underline{v}_H(\underline{p}_H)$ as the unbadged quality that would be obtained following an unbadged price $\underline{p}_H$ and in which all sellers act optimally. Following the logic described for $\underline{p}_H = L$, we can explicitly write the function $\underline{v}_H(\underline{p}_H)$ for any $M > \underline{p}_H > L$ as follows:

$$\underline{v}_H(\underline{p}_H) = \frac{\mu_\ell G(\underline{p}_H)L + \mu_s G(\underline{p}_H)(1 - F(H - \underline{p}_H)M}{\mu_\ell G(\underline{p}_H) + \mu_s G(\underline{p}_H)(1 - F(H - \underline{p}_H)}$$

As established above, $\underline{v}_H(L) > L$, and by the same logic, $\underline{v}_H(M) < M$ because both shirking s-types and ℓ-types will enter and be unbadged. Because both $G(\cdot)$ and $F(\cdot)$ are continuous, the function $\underline{v}_H(\underline{p}_H)$ is continuous, and must cross the 45-degrees line at least once. This proves than an equilibrium exists.

$\square$

The left panel of Figure 1 illustrates the logic of Proposition 1. The upshot from the description of equilibria above is that any equilibrium $B = H$ will satisfy the following:[4]

1. **prices:** $\bar{p}_H = \bar{v}_M = H$, and $\underline{p}_H = \underline{v}_H \in (L, M)$,
2. **entry:** $\mu_{\ell H} = G(\underline{v}_H)\mu_\ell$, $\mu_{sH} = G(\underline{v}_H)\mu_s + \int_{\underline{p}_H}^{H} \int_{H-\underline{p}_H}^{H-k} dG(x)dF(y)dxdy$, and $\mu_{hH} = G(H)\mu_H$,
3. **behavior:** Some s-types who enter choose to work and some to shirk. The measure of s-types who shirk is $G(\underline{v}_H)(1 - F(H - \underline{p}_H))\mu_s$

Note that there may potentially be more than one equilibrium, and conditions on $G(\cdot)$ and $F(\cdot)$ can be described to guarantee uniqueness, yet this is not a concern given our interest in comparing any equilibrium with $B = H$ to the unique equilibrium with $B = M$.

[4] The double integral represents the s-types who enter with $\underline{p}_H < k < H$ and for whom $e + k < H$ so they prefer to enter and work over exiting or entering and shirking.

3.3 Comparative Statics

We now compare what happens when the badging requirement becomes more stringent, and changes from $B = M$ to $B = H$. The following four corollaries follow immediately from comparing prices across the two equilibria.

Corollary 1. $\underline{p}_H < \bar{p}_M$.

This means that those s-types who lose there badge are hurt by facing a lower price, and hence some of them will exit. This also implies that if there is some natural entry and exit, there will be less entry by s-types.

Corollary 2. $\underline{p}_H > \underline{p}_M$ *and* $\bar{p}_H > \bar{p}_M$.

This implies that both for ℓ- and h-types, prices will be higher with a more stringent badging requirement, and hence, more entry occurs at the tails of the quality distribution. Furthermore, recall that some s-types will now exert effort, implying more entry at quality level H. Together with Corollary 1 this implies that the distribution of entrants will have "fatter tails" after the more stringent badge is implemented.

Corollary 3. *s-types who retain their badge will increase quality and produce H instead of M.*

This follows immediately from the fact all s-types choose to shirk when $B = M$ while some positive measure choose to work when $B = H$. Hence, those sellers who forst lose their badge and then regain it will necessarily have produced higher quality.

Corollary 4. *Let market A have measure μ_s^A and let market B have measure $\mu_s^B > \mu_s^A$, fixing the other measures of ℓ- and h-types across the markets. If both markets experience a change of badge from lax to stringent, then more entry of h-types will occur in market B*

This result follows from the fact that, fixing the measure of h-types, an increase in s-types means a lower price $\bar{p}_M$ in market B. This in turn implies that when the badge becomes stringent, and $\bar{p}_H = H$ in both markets, the badged-price increases more in market B, and hence there will be more entry of both h-types, as well as s-types who choose to work.

4 Data

We use proprietary data from eBay that include detailed characteristics on product attributes, listing features, buyer history, and seller feedback and reputation. We begin with data from October

2008 to September 2010, which include all listing and transaction data in the year before and the year after the policy change.

One important feature of our data is information on product subcategories cataloged by eBay. There are about 400 subcategories, such as Office, Lamps and Lighting, Beads and Jewelry Making, Video Game Memorabilia, Digital Cameras, Makeup, and many others. A subcategory is the finest level of eBay's catalog that includes all listings on the site.[5]

In general, it is hard to observe a firm's entry date before it has made a sale or reached a certain size. In our detailed data, however, we observe when a seller publishes its first listing in different subcategories on eBay. We treat this date as a seller's entry date into the subcategory (which we also refer to as market). Additionally, we observe the number of incumbents in any month, which allows us to compute a normalized number of entrants across subcategories, which we call the entrant ratio.

Finally, the use of internal data allows us to construct a measure of quality that is not observed with information that appears on the page. Every seller has a reputation score and percent-positive (PP) on eBay, the latter being the number of positive ratings divided by the total number of ratings. Nosko and Tadelis [2015] demonstrate the extreme skewness of PP, where the mean is 99.3% ad the median is 100%. This finding is consistent with previous literature that documents biases in reviews, e.g., Dellarocas and Wood [2008], Luca [2011], and **?**.

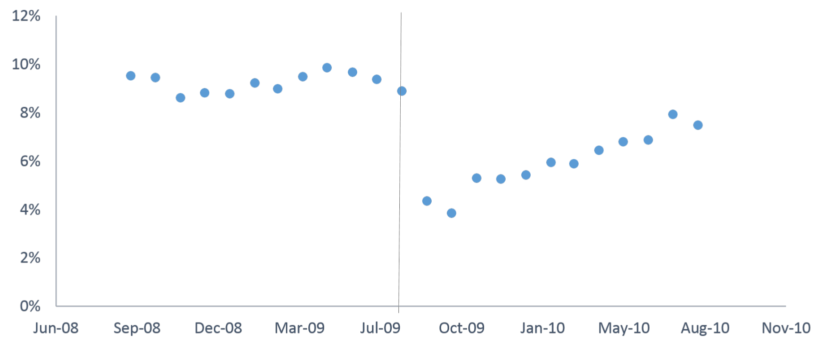
Nosko and Tadelis [2015] conjecture that silence is itself a sign of some negativity, and they construct a new measure that they call "Effective Percentage Positive (EPP), which is the number of positive feedback transactions divided by the number of total transactions. Nosko and Tadelis [2015] have shown that EPP contains much more information on transaction quality than conventional feedback and reputation scores. We follow their approach and for each seller we compute its EPP and use it as a measure of quality.

5 Empirical Strategy

We use the policy change described in section 2 as a quasi-experiment. Figure 2 clearly shows the policy change caused a significant decrease in the share and number of badged sellers. The average share of badged sellers was around 10% throughout the year before the policy change, and dropped to 4% right after the policy change, with some adjustment taking place in the following year.

[5]Prior work has used product ID for finer cataloging (Hui et al. [2016] and Hui [2017]), but these product IDs are only defined for homogeneous products such as electronics and books.

Figure 2: Share of Badged Sellers



Our goal is to create treatment and control groups using variations in policy exposure across different subcategories on eBay. Consider two subcategories on eBay; after the policy change, one market loses all its badged sellers and the other one loses none. The idea is to use outcome variables in the second market as the counterfactual for the first market had the policy not taken place. Our assumption here is that a market with zero drop in share of badged sellers should not have changes in outcome variables. Also, we assume that this variation is exogenous and will later test this assumption with different measures of policy exposure in the online appendix as well as a placebo test. A similar approached is used in Mian and Sufi [2012].

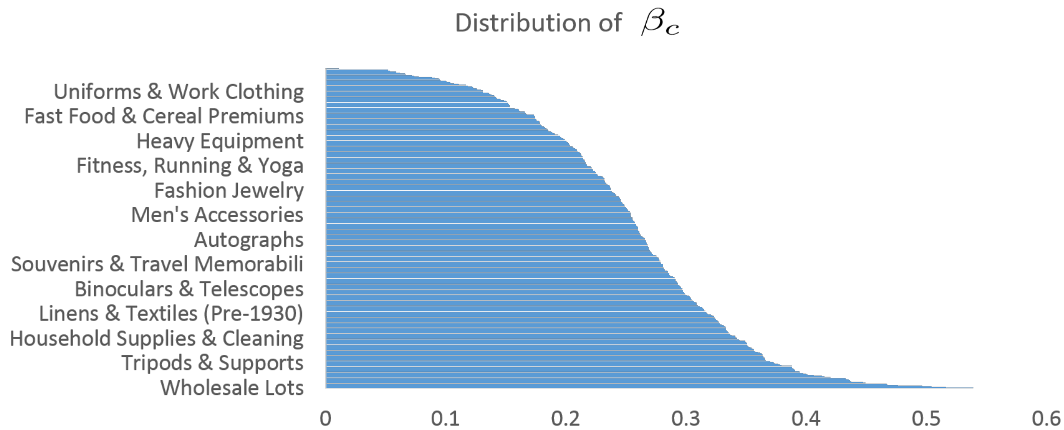
To measure the policy exposure across markets, our first stage is to simulate the percentage drop in the share of badged sellers. In particular, we apply the new certification requirements on badged sellers in the month before the policy change and compute the drop in number of badged sellers divided by the total number of badged sellers.[6]

The horizontal bars in Figure 3 are the ex-ante measure of policy exposure, which is the simulated percentage drop in the share of badged sellers for different subcategories. The figure shows that the decrease in the share of badged sellers after the policy change varies dramatically across different subcategories and for some subcategories is as large as 50%. This large variation in the share of badged sellers after the policy change across subcategories is robust to different specifications discussed above.

Our identification strategy exploits the variability in policy exposure in different subcategories

[6] We establish the robustness of our results by using other measures of the policy exposure and report the results in the online appendix. In particular, we tried 1) immediate change in share of badged sellers using data from the week before and the week after the policy change, 2) estimating the change using an event study in the one, three, and six months before the policy change, 3) using the drop in number of badged sellers instead of shares, and 4) using the percentiles of measures of policy exposure across subcategories. Note that our preferred measure is based on the simulation approach because it is an ex-ante measure of the policy exposure. In particular, in the event study approaches, the change is estimated based on the share of badged sellers after the policy, which itself depends on changes in entry due to the policy change.

Figure 3: Policy Exposure in Different Subcategories



Notes: Policy exposure is the percentage drop in different subcategories on eBay. There are about 400 subcategories, of which the labels on the left are some examples.

induced by the policy change to identify the impact on the number and quality of entrants using a continuous difference-in-difference (DiD) approach. In particular, we estimate the policy impact by comparing the intertemporal changes in the number and quality of entrants in the subcategories that are more affected by the policy change against intertemporal changes of these two measures in the subcategories that are less affected over the same time periods.

This DiD approach is continuous in the sense that the "treatments" (i.e., policy impacts on the share of badged sellers across subcategories) take continuous values between 0 and 1. Specifically, the DiD specification is given as

$$Y_{ct} = \gamma \widehat{\beta}_c Policy + \mu_c + \xi_t + \epsilon_{ct}, \tag{1}$$

where Y_{ct} are the outcome variables of interest in subcategory c in month t (e.g., quality, or entry); $\widehat{\beta}_c$ is the simulated policy impact on the share of badged sellers from our first stage shown in Figure 3; μ_c are subcategory fixed effects; ξ_t are month fixed effects; and ϵ_{ct} are error terms.

Our coefficient of interest is γ, which indicates the percentage change in the outcome variable as a result of variations in the share of badged sellers due to the policy change. Specifically, a statistically significant negative $\widehat{\gamma}$ means that a larger decrease in the share of badged sellers increases the outcome variable, because the signs of $\widehat{\beta}_c$ are negative.

Note that we have two types of entrants: new sellers on eBay (15%) and existing sellers entering new subcategories (85%). An implication of our theoretical framework is that these two types of entry may behave differently if they differ in their entry costs, which is a reasonable assumption. In our main analyses we treat both new sellers on eBay and existing sellers entering new markets of eBay as entry. In Section 7, we repeat our analyses for the two sets of entrants separately and the results are similar.

The DiD approach controls for time-invariant differences in the variables of interest across subcategories; for example, the entrant ratio in the Clothing market is higher than that in the Antiques market. The approach also controls for differences in the entrant ratio over time, for example, changes in the overall popularity of selling on eBay over time. As in most DiD approaches, our key identification assumption for a causal interpretation of $\widehat{\gamma}$ is that serially correlated unobserved errors do not systematically correlate with $\widehat{\beta_c}$ and Y_{ct} simultaneously. We provide a robustness test of this identification assumption in Table 5 in Section 7.

6 Results

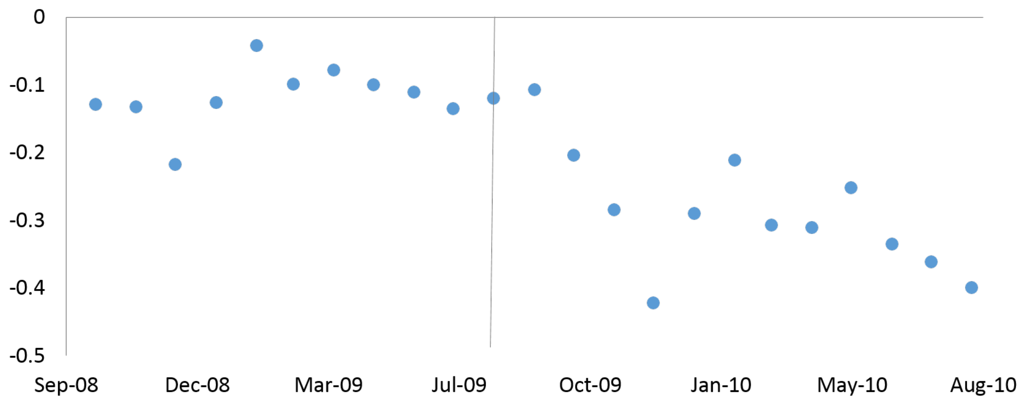
We first estimate the effects of the policy change on the average rate of entry and quality provided by the entrants. We then describe how the quality distribution of the entrants changes. Subsequently, we investigate what can be learned from how incumbents react to the policy change, in comparison to entrants. Finally, we study how prices, sales probability, and market share for different groups of sellers change.

6.1 Effect on Number of Entrants

Figure 3 shows that the policy change resulted in a great deal of heterogeneity in changes in the share of badged sellers across different subcategories. Our theoretical benchmark suggests that variation in how hard it is to obtain a badge will impact both the number and quality of entrants. Before conducting the two-stage regression analysis described above, we first provide descriptive evidence that the heterogeneity in policy impact has meaningful implications for the rate of entry. To normalize across subcategories, we define the entrant ratio to be the number of entrants in month t divided by the number of sellers in month $t-1$ in a particular category. Figure 4 plots the correlation between the entrant ratio and the share of badged sellers in each subcategory.

Given that both the entrant ratio and the share of badged sellers responded to the policy change, we proceed with some simple descriptive facts by normalizing these two measures for a meaningful comparison. In particular, we compute the percentiles of these two measures across subcategories

Figure 4: Market Structure and Entry



Notes: The entrant ratio is defined as the number of entrants in month t divided by the number of sellers in month $t-1$. The percentiles of both variables are defined across subcategories.

and plot their correlation. Figure 4 shows that there is a negative correlation between the entrant ratio and the share of badged sellers across subcategories, i.e., subcategories with a larger share of badged sellers are associated with smaller entrant ratios. This correlation becomes more negative after the policy change, marked by the dashed vertical line, suggesting that the policy change affected the entry pattern in different subcategories, and the magnitude is correlated with changes in the share of badged sellers.

Table 1 reports $\widehat{\gamma}$ from regression (1) for six variables measuring entry, each in a separate panel. Recall that a positive γ means that the increase in the outcome variable is larger in subcategories that have larger policy exposure, i.e. larger drop in share of badged sellers. Panel A column 1 shows that the entrant ratio is higher in subcategories that are more affected, using data from three months before and after the policy change. In particular, a 10% larger decrease in the share of badged sellers leads to 1.2% more entrants. The estimate in column 2 is less negative when we use data from six months before and after the policy change. In column 3, we study the impact seven to twelve months after the policy change,[7] where the estimate is smaller and is not statistically significant. This suggests that the market stabilizes on a new equilibrium after the first six months.

To understand the distributional impact of the policy change on the number of entrants, in Figure 5a we plot two time series, monthly average (normalized) number of entrants and monthly average share of badged sellers, in the subcategories that are most affected (top quintile of β_c) and least affected (bottom quintile of β_c), respectively. Figure 5a shows that in the top quintile

[7]We do not include longer time periods because eBay introduced eBay Buyer Protection in September 2010.

of the subcategories, the share of badged sellers decreases from about 35% to less than 15% right after the policy change, whereas in the bottom 20 percentile, the share of badged sellers decreases from about 15% to 5%. On the other hand, the average monthly number of entrants in the top quintile increases by about 25%, whereas there is no obvious change in the number of entrants for the bottom quintile of subcategories. This suggests that the policy effect on entry comes mainly from subcategories that were heavily affected. Additionally, the entry rates seem to stabilize after three months. We show the robustness of these results by looking at top and bottom deciles of β_c in Figure 9.

6.2 Performance of the Entrant Cohort

We now study how the performance, or quality, of entrants is affected by the policy change using five measures of performance for entrants: EPP, the average entrant size measured by sales quantity, total sales quantity from all entrants, average entrant size in the second year after entry, and total sales quantity from all entrants in the second year after entry. The last two variables are intended to capture the survival rate with a continuous measure.

We construct a seller's EPP using the number of transactions and the number of positive feedback in the first year of entry, conditional on the entrant's survival in the second year (selling at least one item in both the first and second years after entry). The conditioning is intended to eliminate the survival effect from the size effect. We have also tried alternative variations of EPP with different time intervals and without conditioning on survival of sellers; the results are reported in section 7 and show similar patterns.

Negative coefficients in Panel B in Table 1 show that there is an increase in the average quality of entrants in the more affected subcategories after the policy change. For a market with a 10% larger drop in share of badged sellers, the policy effect goes from 0.64% to 0.39% as we expand the window length from six (+/- 3 months) to twelve months (+/- 6 months). Column 3 shows that the increase in EPP persists from the seventh to the twelfth month after the policy change, suggesting that the policy impact on entrants' quality is persistent over a longer time period.

To study the distributional impact, in Figure 5b we show the average EPP for entrants in the top and bottom quintiles of the affected subcategories. Note that EPP is decreasing on eBay over time because buyers are less likely to leave feedback in general, but the average EPP is higher for the top quintile of the affected subcategories compared to the bottom quintile.

Next, we look at the average and total sales quantity during a seller's first year of entry,

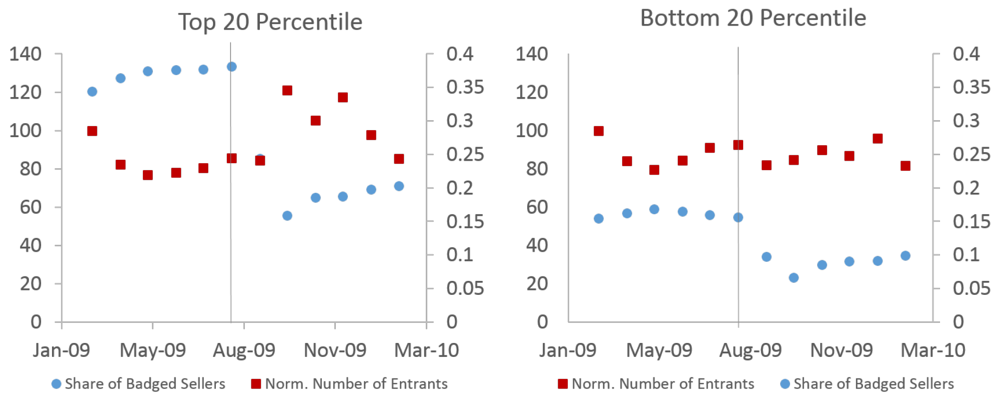
Table 1: Policy Impact on Rate and Quality of Entrants

Panel A. Entrant Ratio			
	(1)	(2)	(3)
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.124***	0.066***	0.010
	(0.021)	(0.016)	(0.032)
R^2	0.911	0.888	0.685
Panel B. EPP Conditional on Survival in the Second Year			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.064***	0.039***	0.043***
	(0.019)	(0.014)	(0.016)
R^2	0.771	0.728	0.699
Panel C. Average Entrant Size Conditional on Survival in the Second Year			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-4.702*	-0.688	-0.628
	(2.563)	(1.792)	(2.084)
R^2	0.612	0.553	0.528
Panel D. Total Sales			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	1326.530*	1090.567	4028.375
	(3908.002)	(2551.776)	(2261.147)
R^2	0.927	0.928	0.943
Panel E. Average Entrant Size in the Second Year			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-1.380	0.341	-1.588
	(2.162)	(1.577)	(1.339)
R^2	0.422	0.354	0.398
Panel F. Total Sales in the Second Year			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	2898.81*	2814.841	-7991.153
	(7018.056)	(4459.005)	(4676.972)
R^2	0.720	0.720	0.723

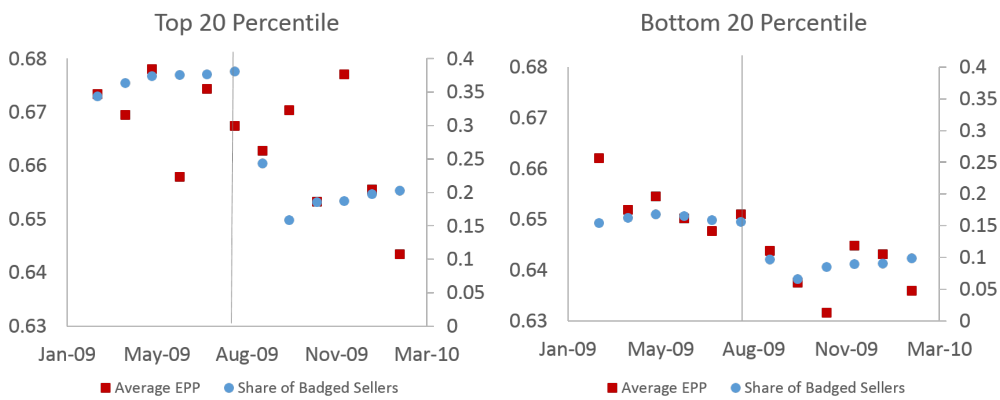
Notes: The regressions are at the subcategory-month levels. An entrant survives the second year if she sells at least one item in both the first and second years after entry.

*** indicates significance at p $\leq$ 0.01; ** p $\leq$ 0.05; * p $\leq$ 0.10.

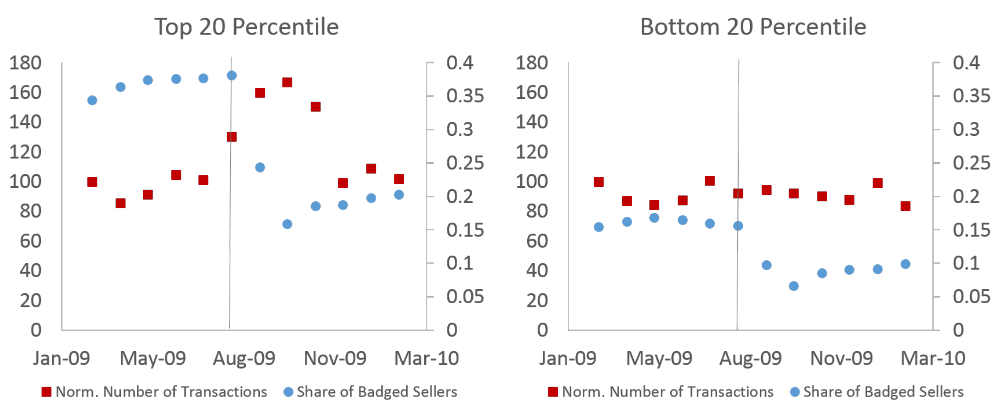
Figure 5: Distributional Policy Impact on Entrants



(a) Distributional Policy Impact on Number of Entrants



(b) Distributional Policy Impact on EPP



(c) Distributional Policy Impact on Sales

Notes: The vertical axis on the right shows the average monthly share of badged sellers, and the one on the left shows the average monthly normalized number of entrants, average monthly EPP, and average normalized number of transactions. The numbers of entrants in the six-month period before the policy change are normalized to 100. The numbers of transactions in the six month months before the policy change are normalized to 100.

conditional on her survival in the second year. A negative coefficient in column 1 of Panel C shows that over the short term, the sales quantity from each entrants is smaller in subcategories affected more by the policy change; however, this drop becomes much smaller and insignificant when considering a longer time period after the policy change. This result indicates that the average entrant is smaller in the subcategories most affected by the policy change. Recall that these subcategories have more entrants on average as well. As a result, this regression does not necessarily imply a decrease in the total number of sales by entrants as a whole. In fact, when we run a regression of total sales from all entrants in Panel D, we observe that the subcategories more affected by the policy change have a higher total number of sales by entrants.

We also plot graphs to analyze the distributional policy impact on entrants' sales by the top and bottom quintiles of the affected subcategories in Figure 5c, which shows a short-run surge in the number of total sales in the top quintile of the affected subcategories with very little impact on the bottom quintile of the affected subcategories.

Finally, we study entrants' survival by looking at the average size of entrants in the year after entry, assigning zero to sellers who do not sell any items in their second year. The advantage of this measure over a simple survival dummy is that it is able to capture a seller's change in size as well as exit.[8] Panel E shows that the average sales quantity in the second year per entrant decreases more for entrants in subcategories that are more affected by the policy change. This observation is consistent with entrants being smaller in the affected area, as shown in Panel C, although none of the estimates are statistically significant. In Panel F, we see that the total number of items sold in the second year by entrants as a whole increased in the short term.

6.3 Quality Distribution of the Entrant Cohort

Corollaries 1 and 2 of our theoretical framework predict that there should be more entrants of high quality, as the benefit of getting a more selective and informative badge is higher. Additionally, the theory predicts that low-quality sellers will enter more often because they are pooling with a better set of sellers who lost their badge, implying higher average prices and/or sales for unbadged sellers in equilibrium. To test this prediction, within each subcategory, we partition entrants into deciles based on their EPP score in the first year after their entry. For example, we look at entrants within the top decile as determined by their EPP score based on their transactions in the first year

[8] Another method to study the survival rate is to have a dummy variable equal to zero if the seller does not sell any item in the second year. However, this is not an appealing measure, as many sellers, even if they quit selling professionally on eBay, may still sell occasionally on the platform.

after their entry. Then we perform the DiD specification for this decile and check if these EPPs have increased more for the subcategories more affected by the policy change. When we look at the top decile (entrants with highest quality in the distribution), a positive number will indicate that average entrant quality is higher in markets with higher policy exposure. This means that we have a fatter tail of the high-end of the distribution. Respectively, if we look at the bottom decile (entrants with lowest quality in the distribution), a negative estimate will mean that average entrant quality is lower in markets with higher policy exposure. This indicates a fatter tail on the low end. Another prediction was that the sellers who had a chance of becoming badged before and no longer have this opportunity after the policy change will enter less often. A distribution of entrants' quality with a fatter tail from both left and right will indicate a smaller share of average-quality entrants.

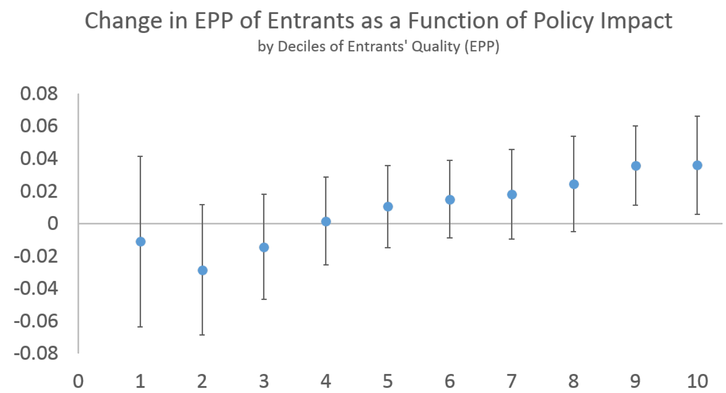
We plot the change in first-year EPP for entrants of different quality deciles in Figure 6. For consistency, we condition the EPP calculation on an entrant's survival in the second year. Entrants are counted every two months. To be able to take the average of cohorts, we restrict our attention to subcategories with at least 100 entrants. As a result, for each subcategory, we have three observations (six-month equivalent) before the policy change and three observations after it. Additionally, we only consider subcategories that have entry in all of the six two-month periods and remove subcategories with a small number of entrants. This leaves us with 228 out of the 400 subcategories.

The x-axis in Figure 6 indicates different quality deciles, with "10" being the highest decile of EPP and "1" being the lowest decile of EPP. The figure plots point estimates of the changes in EPP for the entrant cohorts with 95% confidence intervals. The top-two decile point estimates are statistically positive, as predicted. Though the other estimates are not significant from zero, we do observe a monotonically increasing relationship that is consistent with our prediction that the quality distribution of entrants after the policy change varies and has fatter tails because sellers from the extremes of the quality distribution now have stronger incentives to enter. This in turn implies that sellers in the middle of the quality distribution enter less frequently.

6.4 Impact on Incumbents

Figure 7 plots the average monthly EPP of incumbents and entrants in the six months before and after the policy change. The x-axis is the normalized month relative the when the policy change took place. Incumbents are defined as sellers who listed at least one item before and one item after the change. The EPPs for incumbents are computed using transaction in a given month to capture

Figure 6: Change in EPP for Entrants in Different Quality Deciles



Notes: Bars indicate 95% confidence intervals.

potential changes in behavior in that month. Entrants in a month are those who have their first listing in that month. The EPPs for entrants are defined based on the transactions in the year after their first listing.

The blue series show that their is an increase in entrants' EPP, which is consistent with our previous results in Table 1. On the other hand, there seems to be a break in trend for the red series after the policy change. We need to be cautious in interpreting this result, because the change could be due to seasonality, e.g., buyers could be more likely to leave feedback from September to February than the other half of the year.

To deal with this concern, we perform the DiD regression on incumbents in Table 2. In Panel A, we see that although the policy change seems to increase incumbents' EPP in markets with higher exposure to the policy, the changes are not statistically significant at the 10% level. This exercise shows that EPP of incumbents did not increase significantly after the policy change, although a positive estimate may suggest that some groups of incumbents increase their quality.

Next, we study whether there is any change in quality from some groups of incumbents. We first repeat the DiD analyses for sellers who entered not too early before the policy change. The idea is that these sellers could be similar to those that entered right after the policy change because of their proximity in entry date. In Panel B of Table 2, we study how EPP changes for sellers that entered either three months or six months before the policy change. The insignificant estimates show that there are little changes in behavior for these two groups of sellers, suggesting that a significant share of the changes in EPP from entrants is likely to come from improved selection.

Recall that Corollary 3 states that those incumbents who would lose their badge would have to

Figure 7: Change in EPP of Incumbents and Entrants

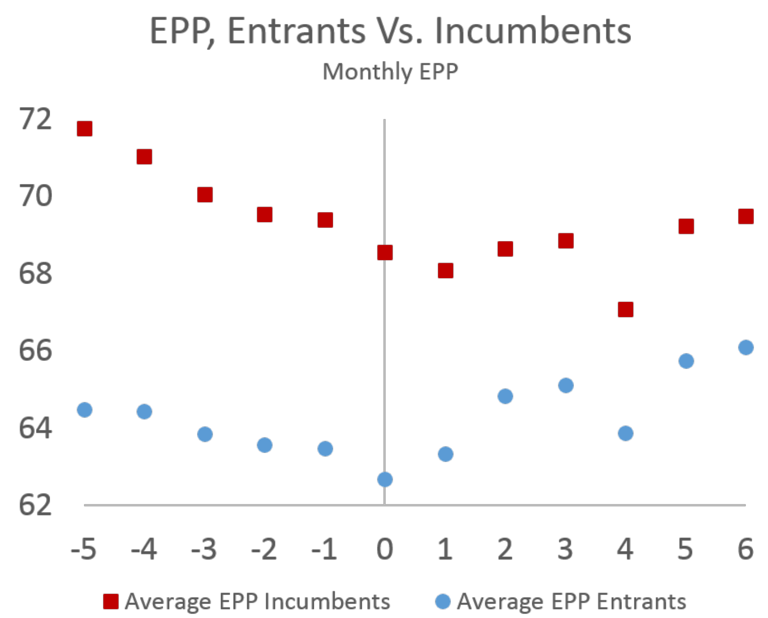


Table 2: Policy Impact on Quality of Incumbents

	(1)	(2)	(3)
Panel A. EPP from Incumbents			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.023	0.019	-0.012
	(0.015)	(0.012)	(0.013)
R^2	0.899	0.869	0.860
Panel B. Sellers who Entered n Months before the Policy			
	$n = 3$	$n = 6$	
Estimate	-0.042	-0.058	
	(0.027)	(0.050)	
R^2	0.463	0.409	

Notes: The regressions are at the subcategory-month levels. An incumbent is defined as a seller who has listed at least one item before and one item after the policy change in the specified time windows. *** indicates significance at $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.10$.

exert effort to increase their quality if they wish to maintain their badge. Because incumbents can be badged or not badged before and after the policy change, we divide them into four collectively exhaustive groups based on their certification status before and after the policy change. One is the group of sellers who were badged both before and after the policy change, which we denote group BB. Another consists of sellers who were badged before the change but had no badge after, which we denote BN. We similarly define groups NB and NN. A seller was badged before the policy if she was badged for at least five out of six months before the policy change.[9] The seller's badge status afterwards depends on whether she meets the new policy requirements by the end of the day before the policy change. In other words, a seller's badge status before the policy is based on the actual measure and her status after is based on simulation. In the online appendix, we also repeat the analyses using seller's actual status after the policy change to define the four groups, and the results are similar. The largest group is the NN group with over 50% of sellers, while the NB group is the smallest at 4%.

We perform the DiD analyses on the four groups of incumbents in Table 3. In Panels A–D, wee see that there is no statistically significant change in incumbents' quality when we look at the sample period from three months before and after the policy change.[10] When we look at six months before and after the policy change, the only group that experience a larger increase in EPP in more affected market is group BN. This result is consistent with Corollary 3: some sellers who lost their badge due to the new policy will increase their quality to meet the new badge requirements.

To analyze this further, we study the change in the behavior of BN incumbents based on whether they regain their badge within the three months after the policy change. We see in Panel E that, a BN incumbent who regained her badge in the near future increases her quality in the three and six months after the policy change. On the other hand, an BN incumbent who remained unbadged in the near future does not increase their quality in neither the three months or six months before the policy change. This is consistent with Corollary 3.

To sum up, we find that incumbents do not increase in quality on average except for a particular group, which are likely to be those marginal incumbents who are badged before the policy change, lose their badge after the change for for a short while until they regain their badges.

Disentangling improved selection from better behavior is tricky for entrants. The reason is that,

[9]For robustness, we also change the threshold for each group to three and four months out of six. The results are qualitatively similar.

[10]In the NN group, we only look at incumbents who have sold at least 6 items in the 6 months before the policy change to get rid of occasional sellers.

Table 3: Policy Impact on Different Incumbent Groups

	+/- 3 Months	+/- 6 Months	Month 7 to 12
Panel A. BB Incumbents			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.067	0.048	0.107***
	(0.047)	(0.039)	(0.041)
R^2	0.661	0.534	0.509
Panel B. BN Incumbents			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-0.018	0.043**	0.086***
	(0.028)	(0.020)	(0.023)
R^2	0.820	0.779	0.753
Panel C. NB Incumbents			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-0.064	0.014	-0.001
	(0.059)	(0.041)	(0.044)
R^2	0.494	0.473	0.474
Panel D. NN Incumbents			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-0.012	0.007	0.051
	(0.038)	(0.028)	(0.031)
R^2	0.692	0.648	0.624
Panel E. BN Incumbents who Regain Badge in 3 Months			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.084**	0.121***	0.134***
	(0.041)	(0.032)	(0.035)
R^2	0.705	0.610	0.590
Panel F. BN Incumbents who Remain Unbadged in 3 Months			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-0.044	0.005	0.051**
	(0.029)	(0.022)	(0.024)
R^2	0.783	0.740	0.720

Notes: The regressions are at the subcategory-month levels. Badge statuses are simulated by applying the new policy requirements on incumbent sellers. An incumbent is defined as sellers who list at least one item both before and after the policy change.

*** indicates significance at $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.10$.

unlike in the case of incumbents, we cannot fix a set of entrants and track their behavior before they enter the market. However, because the change in incumbents' behavior overall is small, and is attributed only to a small number of incumbents, we believe that a significant fraction of the increase in quality by entrants at the tails of the quality distribution is likely due to selection rather than to behavioral changes. On the other hand, entrants whose quality level are similar to the incumbent group in Panel E in Table 3 may increase their quality after the policy change.

6.5 Impact on Badge Premium

After the policy change, consumers will see fewer badged sellers in the search result page. This should in theory change their valuation towards the badge, and could change the price and sales probability of sellers of different types either because consumers understand the higher quality threshold, or just the simple fact that demand for badged sellers now faces a smaller supply. In this section, we study how badge premiums change for the four groups of sellers (BB, BN, NB, NN) defined previously.

Following the literature that studies price changes on eBay (e.g., Einav et al. [2011] and Hui et al. [2016]), we take advantage of product ID's in our data to construct an average price for each product that was listed as fixed-price and sold. For each individual item sold we define its "relative price" as the item's price divided by the average price of the product. In column 1 of table 4, we study the change in the price premium, which is the relative prices, for different groups of sellers using transactions from one month before and one month after the policy change, where NN is the excluded group. We find that the sellers in the BN group experience a statistically significant decrease in relative price of 5.2% (relative to the 1.5% decrease in the NN group). The changes in relative price for the other groups are not statistically different from the change in the NN group. We should note that one new benefit of the eTRS badge is a 20% discount in the commission fee, which is like a tax reduction on revenue for sellers in the BB and NB groups. The average commission rate on eBay is 15%, and therefore a 20% reduction is equivalent to a 3% price increase. Some of this benefit may be passed through to buyers due to competition on the platform.[11]

In columns 3 and 4, we show the changes in badge premium in terms of sales probability and sales quantity using transactions from one month before and one month after the policy change. We see that all groups of sellers except for group BN experience an increase in both measures. The magnitudes for both measures in descending order are NB, BB, NN, and BN. Our interpretation

[11] Our theoretical model ignores pass-through by assuming that consumers pay their willingness-to-pay.

Table 4: Change in Badge Premium

	(1) Price	(2) Relative Price	(3) Sales Probability	(4) Sales Quantity	(5) Market Share
Policy	-0.035	-0.015***	0.015***	0.010**	3.6E-06(21%)*
	(0.029)	(0.003)	(0.001)	(0.005)	(2.2E-06)
BB*Policy	-0.258	0.010	0.032***	0.035***	3.9E-06(5%)***
	(0.167)	(0.014)	(0.001)	(0.012)	(5.3E-06)
BN*Policy	-0.517***	-0.052***	-0.014***	-0.031***	-1.1E-05(-16%)
	(0.074)	(0.006)	(0.001)	(0.004)	(2.7E-06)
NB*Policy	0.421	-0.057	0.041***	0.149***	-3.4E-07(-1%)
	(0.456)	(0.041)	(0.002)	(0.015)	(7.2E-06)
Seller FE	✓	✓	✓	✓	✓
Product FE	✓	✓	✓	✓	
Week FE	✓	✓	✓	✓	✓
R^2	0.993	0.105	0.845	0.984	0.817

Notes: In columns 1–3, we use transaction data from one month before and one month after the policy change. In columns 2 and 3, we also control for relative price. B (or N) indicates that the seller is badged (or not badged). The first (second) letter refers to the seller's status before (after) the policy change. In column 4, we fill in zero market shares if a seller does not sell in a particular week.

*** indicates significance at $p \leq 0.01$; ** indicates $p \leq 0.05$; * indicates $p \leq 0.1$.

is that the sellers in the NB group experience an increase in sales because they gain the reputation badge.[12] The sellers in the BB group experience an increase in sales because the new badge conveys more information and therefore is more valuable than the old one. The sellers in the NN group are better off because they are being pooled with higher-quality sellers than before. Finally, the sellers in the BN group are worse off because they lose their badge. Combining the estimates for the BN group in columns 2 and 4, we see that they receive a lower price and sell less after the policy change, implying that they are worse off after the policy change. This is consistent with our theory that middle-quality sellers are hurt by the policy change.

Finally, we analyze the policy impact on market share for different groups of sellers using their transactions from one month before and one month after the policy change. This regression is at the seller–week level so that the market share of a seller in a given week equals the number of transactions of that seller divided by the total number of transactions in that week. If a seller does not have any sale in a particular week, we fill in zero as that seller's market share for that week. We report the estimates in column 5 as a percentage of the average market share for the corresponding seller group before the policy change. We see that the BB group experienced an increase in their

[12] Note that this tiny group exists because the certification happens every month. Therefore, an eligible seller still needs to wait till the certification date to get badged.

market share of 15% relative to the benchmark NN group. This translates to a net increase of 5% as well because the change in market share for the NN group is small. On the other hand, the BN group had a 16% smaller (relative and net) market share after the policy change, although the result is not as significant. We also performed the same set of regressions using transactions in the three months before and three months after the policy change. The results are similar and reported in the online appendix.

Consider all the estimates in Table 4 together, we see that after the policy change the BN group is worse off and the other three groups are better off mostly through increased sales.

7 Robustness Checks

In this section, we perform several robustness checks to ensure that our empirical results are robust. We first provide evidence that our identification assumption seems reasonable. Next, we show that our results hold for the two types of entrants we discussed in early, namely, new sellers to eBay, and experienced sellers who enter a new market (subcategory). Subsequently, we show that our result on no change in incumbents' behavior is robust regardless of the time windows used in the definition of EPP. Finally, we provide robustness checks on changes in badge premiums for different groups of incumbents by changing the window size of the estimation. In the online appendix, we show that all of our results are robust to different kinds of first-stage specification, namely using an event-study approach and use a normalized rank-preserving measure of β_c instead using the absolute values.

7.1 Placebo Test on the Exclusion Restriction

Our identification assumption in the difference-in-difference estimation is that there are no serially correlated heterogeneities across subcategories that simultaneously affect both changes in share of badged sellers and changes in entry variables. Like in any other exclusion restrictions, we cannot directly test this assumption. Therefore, we provide some suggestive evidence that the identification assumption does not seem to be violated.

Our thought experiment is the following. Suppose there exist serially correlated category-specific confounders that drive our results, and assume that there is some persistency in this confounding effect over time. This assumption would imply that the estimated change in share of badged sellers in the year of the policy change, which partially stems from the persistent confounding effect, should be able to explain differences in entry patterns in the year prior to the policy change.

We test this using a placebo test: we use the simulated $\widehat{\beta}_c$, and re-perform the second-stage regression using data around September in the previous year. In Table 5, we report the estimated γ for entrant ratio, EPP, and total sales for entrants in the previous year. Neither of the estimates is statistically significant in this table, implying that the impact of the policy change on the share of badged sellers in different subcategories is as good as random with respect to different entry variables across subcategories in the previous year. This suggests that the policy change generates some exogenous variations in share of badged sellers across subcategories that are not mere artifacts of heterogeneities across subcategories. We also repeat the placebo test in the three months and the six months before the policy change, respectively. The estimates are also not statistically significant.

In principle, there could still exist serially correlated confounders that are not persistent over time, and they will contaminate our causal interpretation. Like in any two-step estimation, whether the exclusion restriction assumption holds is an empirical question. However, we believe that the estimates in the placebo test being very noisy is reassuring; for example, the standard error for change in entrant ratio using data from three months before and after is more than four times larger than the point estimate.

7.2 Two Types of Entry

We distinguish between two types of entrants into a subcategory: new sellers on eBay and existing sellers entering a new subcategory. From the lens of our theoretical model, these two types of entrants differ in their entry cost: the cost of entering eBay is higher than the cost of entering a new subcategory for existing eBay sellers. The former requires sellers to understand marketplace, its rules and regulations, and also to decide which items to sell and how to acquire those items to sell. On the other hand, the latter will need to only acquire new items to sell in a new subcategory which may mean finding new resources for their inventory. Differences in fixed cost of entry will result in differences in entry decision of the firms as a result of a change in reputation mechanism.

We find that among entrants into new markets, about 15% of them are new sellers on the platform and 85% are existing sellers entering new subcategories. Next, we perform our previous DiD analyses for the two types separately (see Table 6). Figure 8 shows the change in EPP deciles similar to Figure 6. Results in both exercises are very similar across the two types, and the relative magnitudes of these estimates are consistent with our theory. Assuming that entry costs of starting

Table 5: Placebo Test on the Exclusion Restriction Assumption

Panel A: One Year Before the Policy Change				
	(1)	(2)	(3)	(4)
	Entrant Ratio	EPP	Seller Size	Total Sales
Estimate	0.752	-0.005	1.383	8937.404
	(2.689)	(0.020)	(2.871)	(22757.260)
R^2	0.216	0.718	0.619	0.948
Panel B: Six Months Before the Policy Change				
	(1)	(2)	(3)	(4)
	Entrant Ratio	EPP	Seller Size	Total Sales
Estimate	0.030	0.014	-0.500	-13029.790
	(0.031)	(0.020)	(2.210)	(19913.113)
R^2	0.813	0.729	0.637	0.955
Panel C: Three Months Before the Policy Change				
	(1)	(2)	(3)	(4)
	Entrant Ratio	EPP	Seller Size	Total Sales
Estimate	-0.044	0.025	-4.510	-2139.507
	(0.057)	(0.018)	(4.546)	(32882.516)
R^2	0.839	0.763	0.625	0.959

Notes: We use the $\widehat{\beta}$ estimated from the year of the policy change, and re-perform the second-stage regression using data from both three months and six months before and after September in the previous year.

Table 6: Two Types of Entry

	New Sellers		Existing Sellers	
Panel A. Entrant Ratio				
	(1)	(2)	(3)	(4)
	+/- 3 Months	+/- 6 Months	+/- 3 Months	+/- 6 Months
Estimate	0.033***	0.019***	0.124***	0.066***
	(0.006)	(0.004)	(0.021)	(0.016)
R^2	0.898	0.886	0.911	0.889
Panel B. EPP				
	(1)	(2)	(3)	(4)
	+/- 3 Months	+/- 6 Months	+/- 3 Months	+/- 6 Months
Estimate	0.077*	0.188***	0.043***	0.068***
	(0.043)	(0.059)	(0.014)	(0.019)
R^2	0.298	0.401	0.717	0.746

*** indicates significance at p $\leq$ 0.01; ** p $\leq$ 0.05; * p $\leq$ 0.10.

to sell on eBay are higher than those of entering a new category for an existing seller on eBay, new sellers need to have higher quality to compensate for the entry cost relative to the increase in quality among existing sellers. By the same logic, there should be more entry of the existing sellers relative to the increase in entry of new sellers.

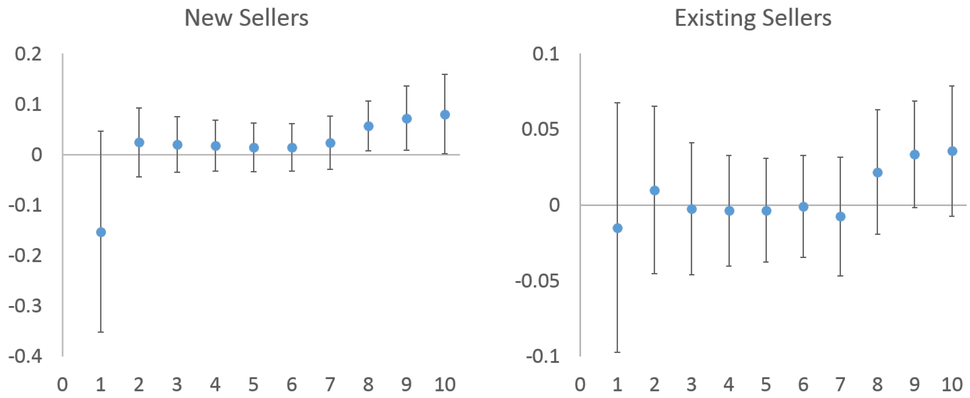
In practice, it is possible that sellers do not make entry decisions into different subcategories based on changes in price in that category alone. Rather, they may have reasons to sell in more than one category, such as meeting both the value and quantity requirement for the eTRS badge that is not category-specific. For example, a laptop seller might find it harder to meet the minimum quantity requirement for the new badge, and therefore enters a cheap category like cables merely to meet this requirement.

Finally, we try to understand the "transition" between existing and new subcategories that sellers operate in, and how this transition varies based on a seller's badge status before and after the policy change. Consistent with our previous results, sellers are more likely to enter markets that are affected more by the policy, but the transition probabilities do not differ statistically across the four groups–BN, BB, NB, and NN. This suggests that sellers in the BN group do not have a larger incentive to enter in order to meet the badging requirements.

7.3 Alternative EPP Measure for Incumbents

In this section, we defined EPP for incumbents based on their transactions in the six months before a particular month. In Figure 10, we plot average EPP for incumbents and entrants using the new

Figure 8: Change in EPP for Two Types of Entrants in Different Quality Deciles



Notes: Bars indicate 95% confidence intervals.

definition of EPP. We see an increase in the EPP of entrants and no obvious change in incumbents's EPP.

In Figure 11a, we plot the average monthly EPP for incumbents in the four groups. The solid line is the average monthly EPP provided by incumbents from a particular group in the six months before and after the policy implementation date. The dotted line and dashed-dotted line are the average EPP provided by the same set of incumbents in the same months in the previous year and the following year, respectively. We see that there is no obvious difference between the EPP provided by incumbents in the year of the policy change and the EPP in adjacent years, except that EPPs are getting lower over time.[13] This implies that the change in the average monthly EPP observed in these two figures maybe due to due to seasonality.

We created a similar plot for sellers of different quality quartiles measured by EPP in Figure 11b. The graphs are similarly constructed, and we again note that there is no observable change in incumbents' EPP after the policy change after removing seasonality. Thus, the incumbents do not seem to change their behavior in response to the policy change based on this definition of EPP.

8 Conclusion

Following a policy change on eBay, more demanding criteria and more precise information are used to award a quality-signalling badge to sellers. We use this change to gauge insight into the effects of more stringent certification and reputation measures on entry, which is a hard-to-study relationship.

[13] This is because buyers are less likely to leave feedback over time.

We exploit the differential impact of the policy change (policy exposure) on different subcategories of sellers for identification, and document a negative correlation between the share of badged sellers and the rate of entry across subcategories affected by the policy change. The subcategories that experience a higher reduction in the share of badged sellers because of the policy change have larger entry rates after the policy change. However, this effect is temporary, and tends to disappear once the market adjusts to the new equilibrium, after about six months.

We also find that the distribution of quality provided by entrants has fatter tails after the policy change. This finding is consistent with the prediction of a simple model where a high bar for certification implies that entrants from both extremes of the quality distribution have stronger incentives to enter. We also find a significant increase in the overall quality provided by entrants in the more affected subcategories, as measured by the EPP, an increase that, contrary to that of entry rates, persists even from the seventh to the twelfth month after the policy change. We find no change in the quality provided by incumbents, however, which suggests that a significant part of the observed change in the distribution of quality provided by entrants is indeed likely to be linked to selection rather than to a change in entrants' behavior. These results indicate that the availability and precision of past performance information are important not only for the rate of entry in a market, but also for the quality of who is actually entering, hence for how markets evolve in the long run.

Our results have implications for the design of reputation and certification mechanisms in digital platforms and other markets with asymmetric information: this design could have significant effect on the number and quality of entrants. The ability to encourage the entry of high-quality sellers is not only important to customers satisfaction from the platforms, but could also be important to innovation in the economy.

References

George A Akerlof. The market for" lemons": Quality uncertainty and the market mechanism. *The quarterly journal of economics*, pages 488–500, 1970.

Patrick Bajari and Ali Hortacsu. Economic insights from internet auctions. *Journal of Economic Literature*, 42(2):457–486, 2004.

Jeffrey V Butler, Enrica Carbone, Pierluigi Conzo, and Giancarlo Spagnolo. Reputation and entry in procurement. 2013.

Judith A Chevalier and Dina Mayzlin. The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3):345–354, 2006.

Pradeep K Chintagunta, Shyam Gopinath, and Sriram Venkataraman. The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, 29(5):944–957, 2010.

C. Dellarocas and C.A. Wood. The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Science*, 54(3):460–476, 2008. ISSN 0025-1909.

Chrysanthos Dellarocas, Federico Dini, and Giancarlo Spagnolo. Designing reputation (feedback) mechanisms. 2006.

Douglas W. Diamond. Reputation acquisition in debt markets. *Journal of Political Economy*, 97 (4):828–862, 1989.

David Dranove and Ginger Zhe Jin. Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, 48(4):935–963, 2010.

Liran Einav, Theresa Kuchler, Jonathan D Levin, and Neel Sundaresan. Learning from seller experiments in online markets. Technical report, National Bureau of Economic Research, 2011.

Daniel W Elfenbein, Raymond Fisman, and Brian McManus. Market structure, reputation, and the value of quality certification. *American Economic Journal: Microeconomics*, 7(4):83–108, 2015.

Gene M Grossman and Henrik Horn. Infant-industry protection reconsidered: the case of informational barriers to entry. *The Quarterly Journal of Economics*, 103(4):767–787, 1988.

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

Xiang Hui. E-commerce platforms and international trade: A large-scale field experiment. *Available at SSRN 2663805*, 2017.

Xiang Hui, Maryam Saeedi, Zeqian Shen, and Neel Sundaresan. Reputation and regulations: evidence from ebay. *Management Science*, 62(12):3604–3616, 2016.

Xiang Hui, Maryam Saeedi, and Neel Sundaresan. Adverse selection or moral hazard: An empirical study. *Working paper*, 2017.

Benjamin Klein and Keith B Leffler. The role of market forces in assuring contractual performance. *Journal of political Economy*, 89(4):615–641, 1981.

Tobias J Klein, Christian Lambertz, and Konrad O Stahl. Market transparency, adverse selection, and moral hazard. *Journal of Political Economy*, 124(6):1677–1713, 2016.

Micheal Luca. Reviews, reputation, and revenue: The case of yelp. com. *Harvard Business School Working paper 12*, 16:12–016, 2011.

Atif Mian and Amir Sufi. The effects of fiscal stimulus: Evidence from the 2009 'cash for clunkers' program. *Quarterly Journal of Economics*, 127(3):1107–1142, 2012.

Chris Nosko and Steven Tadelis. The limits of reputation in platform markets: An empirical analysis and field experiment. Technical report, National Bureau of Economic Research, 2015.

Davide Proserpio and Georgios Zervas. Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science*, 2017.

Carl Shapiro. Premiums for high quality products as returns to reputations. *The quarterly journal of economics*, 98(4):659–679, 1983.

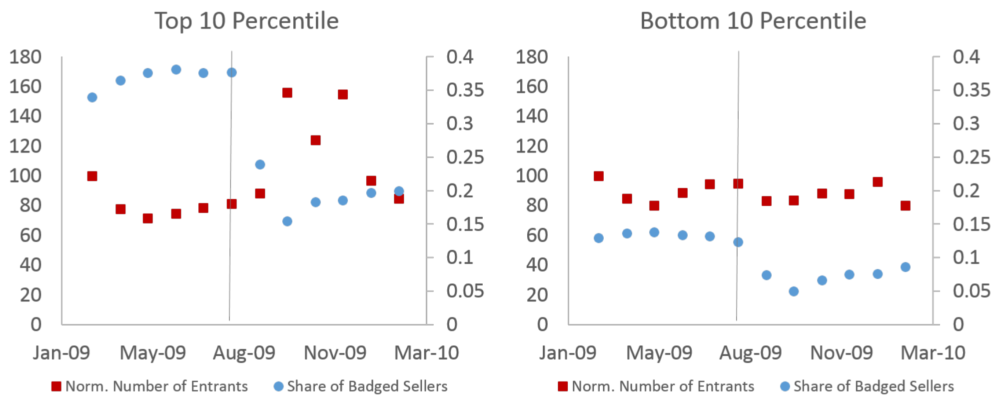
Steven Tadelis. Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8:321–340, 2016.

Chunhua Wu, Hai Che, Tat Y Chan, and Xianghua Lu. The economic value of online reviews. *Marketing Science*, 34(5):739–754, 2015.

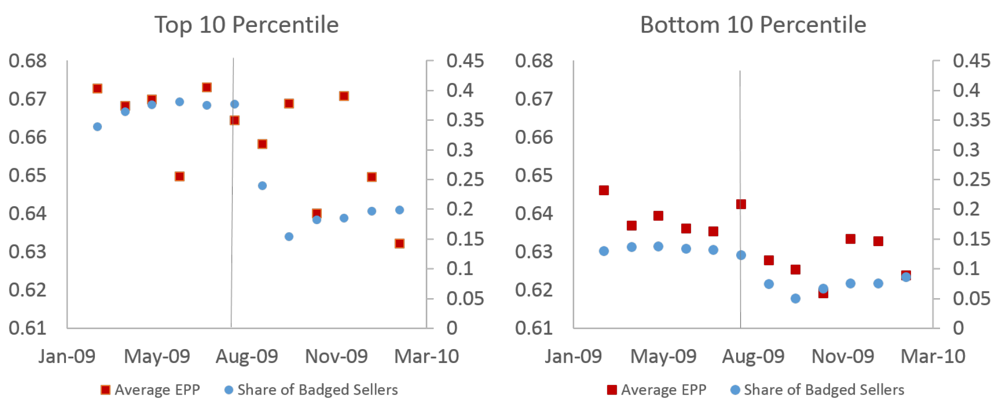
Yi Zhao, Sha Yang, Vishal Narayan, and Ying Zhao. Modeling consumer learning from online product reviews. *Marketing Science*, 32(1):153–169, 2013.

Feng Zhu and Xiaoquan Zhang. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of marketing*, 74(2):133–148, 2010.

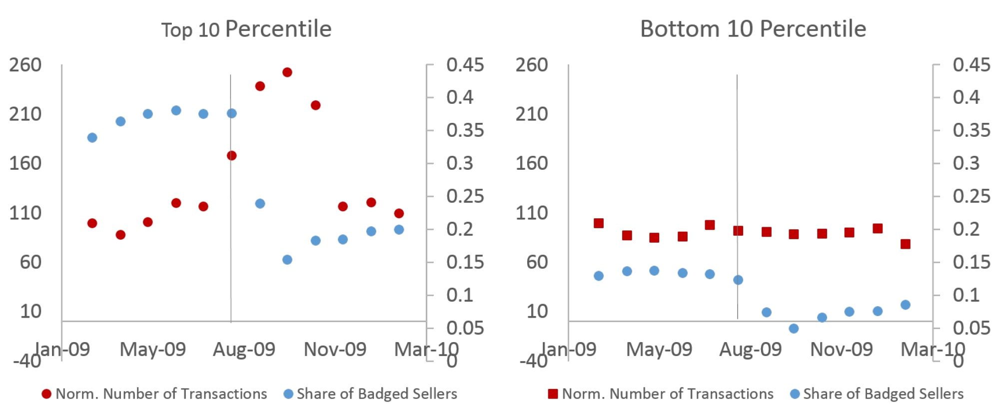
Figure 9: Robustness: Policy Impact on Entrants, Top and Bottom 10 Percentiles



(a) Policy Impact on Number of Entrants



(b) Policy Impact on EPP



(c) Policy Impact on Sales

Notes: The axis for the average monthly share of badged sellers is on the right, and the axis for the average monthly normalized number of entrants, EPP, and the average monthly normalized number of transactions is on the left. The numbers of entrants in the six months before the policy change are normalized to 100. The numbers of transactions in the six months before the policy change are normalized to 100.

Figure 10: Robustness: Change in EPP of Incumbents and Entrants

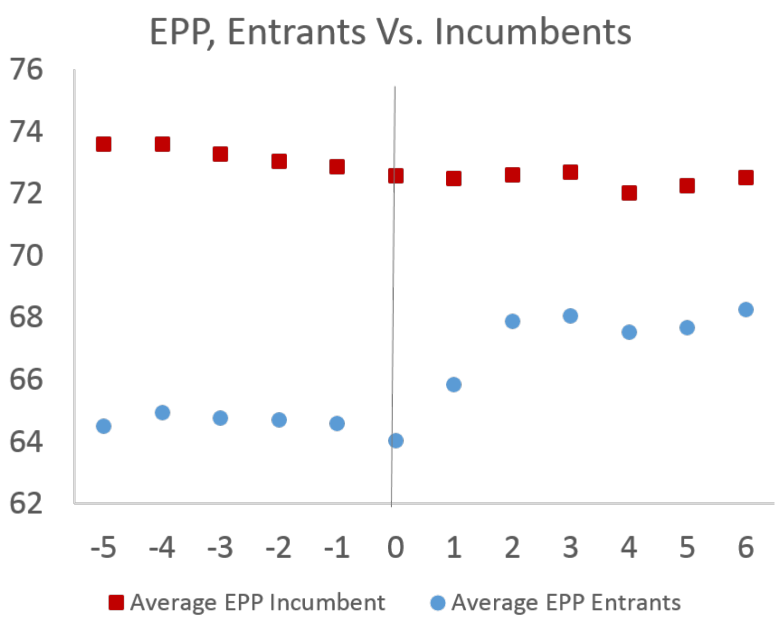
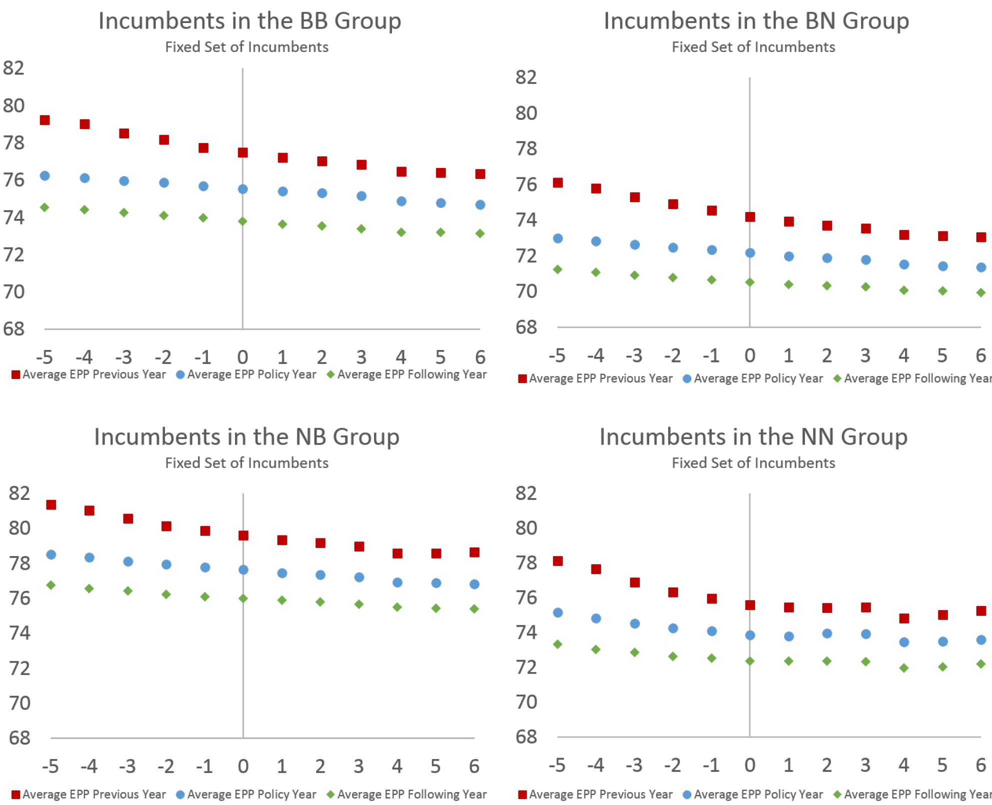
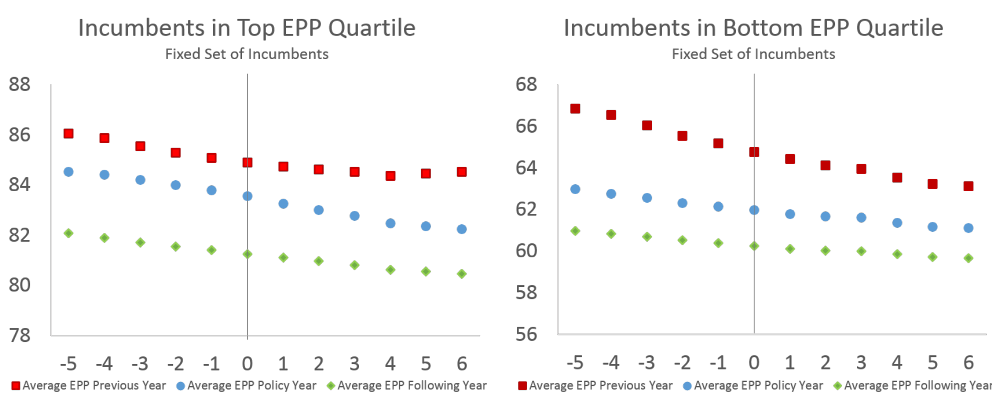


Figure 11: Robustness: Change in EPP of Incumbents



(a) Four Groups of Incumbents



(b) Top Vs. Bottom Quartiles

Notes: The solid line is the average monthly EPP provided by incumbents of a particular group in the year of the policy change. The dotted line and dashed-dotted line are the average EPP provided by the same set of incumbents in the previous year and the following year, respectively. The x-axis shows normalized months, with 0 being the month where the policy change took place.