

# Certification, Reputation and Entry: An Empirical Analysis \*

Xiang Hui <sup>†</sup>	Maryam Saeedi <sup>‡</sup>	Giancarlo Spagnolo <sup>§</sup>	Steve Tadelis <sup>¶</sup>
Washington University	Carnegie Mellon University	SITE, Tor Vergata, EIEF & CEPR	UC Berkeley, NBER & CEPR

January 21, 2019

## Abstract

Markets with asymmetric information will often employ third-party certification labels to distinguish between higher and lower quality transactions, yet little is known about the effects of certification policies on the evolution of markets. How does the stringency in quality certification affect the intensity and composition of entry, incumbents' reactions, and market outcomes? We use detailed administrative data and exploit a policy change on eBay to explore how a more selective certification policy affects entry and behavior across a rich set of online market segments. We find that after the policy change, entry increases and does so more intensely in markets where it is harder to become certified. The average quality of entrants also increases more in the more affected markets, while the quality distribution of entrants exhibits fatter tails *ex post*. Finally, some incumbents increase the quality of their service to maintain certification and deliver higher quality after the policy change. The results help inform the design of certification policies in electronic and other markets with asymmetric information.

*JEL* D47, D82, L15, L86

---

\*We thank Sven Feldmann, Kate Ho, Hugo Hopenhayn, Greg Lewis, Ryan McDevitt, Brian McManus, Peter Newberry, Rob Porter, David Ronayne, Konrad Stahl, and many seminar and conference participants for excellent suggestions, discussions and comments. We are grateful to eBay for providing access to the data and to several eBay executives who provided valuable input.

<sup>†</sup>hui@wustl.edu

<sup>‡</sup>msaeedi@andrew.cmu.edu

<sup>§</sup>spagnologianca@gmail.com

<sup>¶</sup>stadelis@berkeley.edu

# 1 Introduction

Several market institutions have emerged to help mitigate frictions caused by asymmetric information, including warranties ([Grossman \(1981\)](#)), reliance on past reputation ([Shapiro \(1983\)](#)) and regulated certification by a trusted institution ([Leland \(1979\)](#)). 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”.

At the same time that certification badges help alleviate asymmetric information, they can act as barriers to entry for new high-quality entrants who do not have a certifiable track record ([Klein and Leffler \(1981\)](#), [Grossman and Horn \(1988\)](#)). Online marketplaces that choose to use certification badges must therefore understand the ways in which 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 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 marketplace platforms. 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 and was therefore more selective.

Our empirical analyses are guided by a simple theoretical model in which a certification badge impacts the entry and effort incentives of heterogeneous sellers who differ in their quality provision. A more stringent badging requirement causes the average quality of both badged and unbadged sellers to increase because 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 induces more entry of top-end-quality firms by increasing their future payoff from obtaining a more selective badge, but it also induces more entry of low-quality entrants because they will be pooled with better non-badged sellers. Mid-range quality sellers will find entry to be less attractive after the policy change if obtaining a badge becomes too costly, or may change their behavior and exert higher effort if they can profitably earn the more selective badge.

Our theoretical framework generates predictions on how more stringent requirements change outcomes within a single market, as well as across heterogeneous markets differentially. To identify these effects empirically, we exploit differential exposure to the policy change across 400 separate subcategories of the eBay marketplace. Despite the fact that the change in certification requirements is the same across all subcategories, the difficulty to meet the new requirements exogenously differs across subcategories, which in turn leads 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 relative difficulty of obtaining the badge is exogenously different across markets.

Our analysis shows that 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. This appears an ‘adjustment’ effect, as it is significant for the first six months after the policy, after which it fades and becomes insignificant. We then find that the average quality provided by entrants increases significantly after the policy change. In contrast to the effect on the number of entrants, the increase in quality persists over time. Importantly, and consistent with our theoretical framework, we find that the distribution of the quality provided by entrants also changes with the policy and exhibits ‘fatter tails’, in the sense that a larger share of entrants provides quality at the top and bottom quintiles of the quality distribution. When we differentiate new entrants and existing sellers entering a new subcategory, the largest group, we find that results are robust, and that markets more affected by the policy have significantly more entrants with pre-existing certification from other markets, which points at selection as an important driver of changes in quality.

Aside from affecting the selection of entrants, an increase in quality could also be due to sellers changing their behavior and providing higher quality. We therefore study the behavior of four exclusive groups of *incumbent* sellers, depending on whether or not they had a badge before and after the change in policy. Consistent with our model, the only incumbents that show a significant change in behavior are those who lose their badge and, by improving quality provision, manage to re-gain the new badge within three months. These last results confirm 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 likely linked to selection.

We then study how price and market shares changed for these four different groups of incumbent sellers—with and without a badge before and after the policy change. The results we find are intuitive: first, sellers who lost their badge experienced a decrease in the relative price that they receive. Second, sellers who remained badged after the change experienced growth in market share. Third, sellers who were not badged before and after the policy change experience changes that are in between the other two groups.

To ensure that our identification assumption is sensible, we perform two placebo tests as well as a robustness check for identification. An important assumption that we make for the identification step 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. Consider the following thought experiment. Suppose there exist serially correlated subcategory-specific confounders that drive our results, and assume that there is some persistency in this confounding effect over time. This would imply that our measure of policy exposure should be able to explain differences in entry patterns in the year prior to the policy change. In the first placebo test, we test this by using our measure of policy exposure to predict number and quality of entrant in the previous periods. We find no impact, consistent with the exclusion restriction of our econometric specification. Another concern can be that a mean reversion is the real reason we observe more entry into more affected categories. In our second placebo test, we redo our analysis for other time periods before the policy change, when there was no such policy change. We do not find any significant impact. Finally, a concern may be that there might be un-serially correlated confounding factors that are not captured by fixed effects that drive our results and that the first placebo test cannot uncover. To test this to the best of our abilities, we redo our analysis by controlling for various variables that might impact entry or average quality of entrants, such as average feedback ratings of sellers, share of disputes in each category, share of badged sellers, average price in each category, and average size of sellers in each category. The results are mainly unchanged.

Our identification critically depends on our estimates of policy exposure across markets in the first stage. Therefore, we perform two robustness checks to make sure that our results are robust to different specifications for the first stage. First, we estimate changes in share of badged sellers across markets using an event-study approach with different window lengths for this first stage estimation: +/- 6 months, +/- 3months, +/- 4 weeks, +/- 1 week. Second, we adopt an

instrumental variable approach that combines the estimates of policy exposure from the simulation approach and the event-study approach. In particular, we use estimated policy exposure from the event-study approach in the second stage DiD regression, and use simulated policy exposure as the instrument. This is to account for cases where the actual change and the simulated change in share of badged sellers are different. If markets differ in other dimensions, like how easy it is for sellers who lose their badges to quickly regain them, we would expect that impacts on things like entry would be smaller in a market where the predicted reduction in badged sellers did not actually occur. The results are qualitatively similar to the ones using our main specification.

Finally, we explore several other robustness tests. We study how exits have changed and find that the quality distribution of exits has “thinner” tails, which complements the results on entry; and we check the robustness of our results to several alternative econometric specifications (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. \(2018\)](#), which also used eBay data 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 and showed that certification provides more value when the number of certified sellers is low and when markets are more competitive. They did not study the impact of certification on the dynamics of entry and changes in market structure. [Klein et al. \(2016\)](#) and [Hui et al. \(2018\)](#) 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 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. \(2018\)](#) complement [Klein et al. \(2016\)](#) and investigate changes in the size of incumbents. They show that although low-quality sellers do not exit after the policy change, their size shrinks dramatically, accounting for 49%–77% of the quality improvement. In contrast 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 entrants and incumbents before and after the change.

A related literature 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 findings reported in these papers, we found that sellers who were badged both before and after the policy change were of higher quality than sellers who were badged before but not after the change. In addition, sellers who were badged both before and after the policy change also benefited from higher conversion rates because the new badge carried higher value than the old one. Our paper also broadly relates to the literature that ties reputation, certification, and transparency to sales performance, including empirical papers such as Cabral and Hortacsu (2010), Hui et al. (2016) and Fan et al. (2016), and theoretical papers such as Avery et al. (1999), Jullien and Park (2014), Stahl and Strausz (2017), and Hopenhayn and Saeedi (2018).<sup>1</sup> Last, our analyses are related to the empirical literature on adverse selection and moral hazard, e.g., Greenstone et al. (2006), Einav et al. (2013) and Bajari et al. (2014).

Our results help guide 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 where certification is ubiquitous, from financial markets where credit ratings are used to obtain the "investment grade" badge, to many final and intermediate goods markets where labelling institutions certify various forms of quality, to public procurement markets where regulatory certification can significantly change the competitive environment and reduce the costs of public services.<sup>2</sup> According to our findings, if a platform (or a large procurer/buyer) is concerned about too much bunching in the middle of the quality range, while there are two few high and low quality sellers, it should increase the stringency of the certifying badge to stimulate entry at the tails of the quality distribution (and vice versa).

The remainder of the paper is organized as follow. Section 2 provides details about the platform and the policy change while Section 3 analyzes a stylized theoretical model that illustrates how the policy change affects entry and quality choices. Section 4 describes our data and Section 5 discusses our empirical strategy. Our results appear in Section 6, Section 7 deals with endogeneity concerns, Section 8 offers several robustness tests, and Section 9 concludes the paper.

---

<sup>1</sup>See also Bajari and Hortacsu (2004), Dranove and Jin (2010), Cabral (2012), and Tadelis (2016) for surveys.

<sup>2</sup>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, (GAO-12-102R), was rather inconclusive (see further discussions in Butler et al. (2013)).

## 2 Background and Policy Change

eBay is known for its well-studied feedback rating system in which sellers and buyers could rate one another with positive, negative, or neutral feedback. eBay later introduced “detailed seller ratings,” in which buyers give sellers an anonymous rating between 1 and 5 stars along four dimensions (item as described; communication; shipping rate; and shipping speed). To combat concerns that retaliation deters buyers from leaving honest negative feedback, in 2008 eBay made the feedback rating asymmetric so that sellers could only leave positive or no feedback for buyers.

In addition to user-generated feedback, eBay started certifying who 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 the salience of the badged suggesting that the seller is of higher quality.

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 eTRS, a seller must surpass the Powerseller status by *additionally* having at least 100 transactions and selling at least \$3,000 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).<sup>3</sup> The information on dispute rates, only available to 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.

Obtaining the eTRS badge was harder than obtaining the Powerseller badge, but also bestowed 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

---

<sup>3</sup>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. 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.

page, which is the default sorting order, resulting in more sales. Finally, the eTRS badge appears on all the seller’s listings, signaling the seller’s superior quality to all potential buyers.

Note that besides the change in the certification policy, there were two other simultaneous policy changes on eBay.<sup>4</sup> The first one is about easier selling procedures, including faster process for unpaid items, removing buyer feedback if buyer’s dispute has been resolved, easier management of buyers emails, etc. This enhanced selling process is similar across all categories, and should be controlled for by our DiD approach. The second is a change in the search ranking algorithms. The main changes here were: 1) the ranking became based on sales per impression (number of times a listing is shown to customers in the search result page), instead of sales; 2) title’s relevance was enhanced; and 3) top rated sellers were to be promoted in the default search ranking algorithm. For the first two changes in search, we can difference the effects on entry out with our DiD approach. For the last change, we include share of badged sellers in our DiD approach, and still obtain similar results, which are reported in Table 8.<sup>5</sup>

### 3 Certification and Entry: A Simple Model

We present a simple model that incorporates both hidden information (adverse selection) and hidden action (moral hazard) in the spirit of [Diamond \(1989\)](#). The model generates comparative statics that offer a series of testable implications, and clarifies the assumptions needed to empirically identify the effect of a more stringent certification on market outcomes.

**Supply:** Consider a market with a continuum of sellers. Each seller can produce one unit of output with zero marginal costs and fixed costs  $k \in [0, \infty)$ , independently distributed 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  $\mu_\ell$  of “low-quality” sellers, indexed by  $\ell$ , who can only produce low quality  $L$ ; a measure  $\mu_h$  of “high-quality” sellers, indexed by  $h$ , who can only produce high-quality  $H$ ; and a measure  $\mu_s$  of strategic sellers, indexed by  $s$ , who can each choose whether to exert extra effort at a cost  $e$  and produce high quality  $H$ , or whether to shirk at no cost and produce medium quality  $M$ , where  $H > M > L > 0$ . 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

<sup>4</sup><https://pages.ebay.com/co/es-co/sell/July2009Update/faq/index.html#2-1> Accessed on 10/30/2018

<sup>5</sup>Note that ideally we would want to control for number of times a listing has been shown to buyer in the search result page. However as of 2018, eBay only stores that data from 2011.



two dimensions of cost heterogeneity,  $(k, e)$ , while  $\ell$ - and  $h$ -type sellers only differ across  $k$ .<sup>6</sup>

**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.

**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  $\underline{v}_M = L$ .<sup>7</sup>

**Equilibrium:** Let  $\mu_{\theta B}$  denote the measure of type  $\theta$  sellers that enter a market with badge  $B$ . Let  $\pi$  denote the fraction of active  $s$ -type sellers who choose to exert effort. An *equilibrium* for threshold  $B \in \{M, L\}$  consists of (i) a pair of prices  $p_B$  and  $\bar{p}_B$ , (ii) measures of each type of sellers,  $\mu_{\theta B}$ , and (iii) the proportion of  $s$ -type sellers who enter and work,  $\pi$ , 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 switching from  $B = M$  to  $B = H$  so that higher-quality is needed to obtain a badge.

### 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

---

<sup>6</sup>We can alternatively assume that the strategic types will have a baseline level of quality equal to  $L$  instead of  $M$  in absence of exerting any level of effort. Then they can increase their level of quality to  $M$  by paying the cost  $e$ , or increase their level of quality to  $H$  by paying a higher cost, say  $2e$ . The predictions of the model remain mostly the same except for the prediction that prices increase for unbadged sellers. With this extension of the model, the price remains the same for unbadged sellers both before and after the policy change.

<sup>7</sup>Rather than three quality levels and two badge levels we explored a model with a continuum of baseline qualities and the option of increasing quality by exerting effort for all baseline quality types. In that model the badge can be set at any level of quality in the interior of the baseline range. The results are similar though the analysis is more involved and distracts from the key economic forces at play. Namely, by increasing the selectivity of the badge, the distribution of types above and below the badge changes in ways that increase average quality for both groups, and this in turn impacts incentives to work as well as incentives to enter. The model we use is in our view the most parsimonious and easy to follow.

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 characterized 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.

That is,  $\bar{p}_M$  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. The measure of sellers who enter are determined by those who can cover their fixed costs given the two equilibrium prices.

### 3.2 Stringent Badge: $B = H$

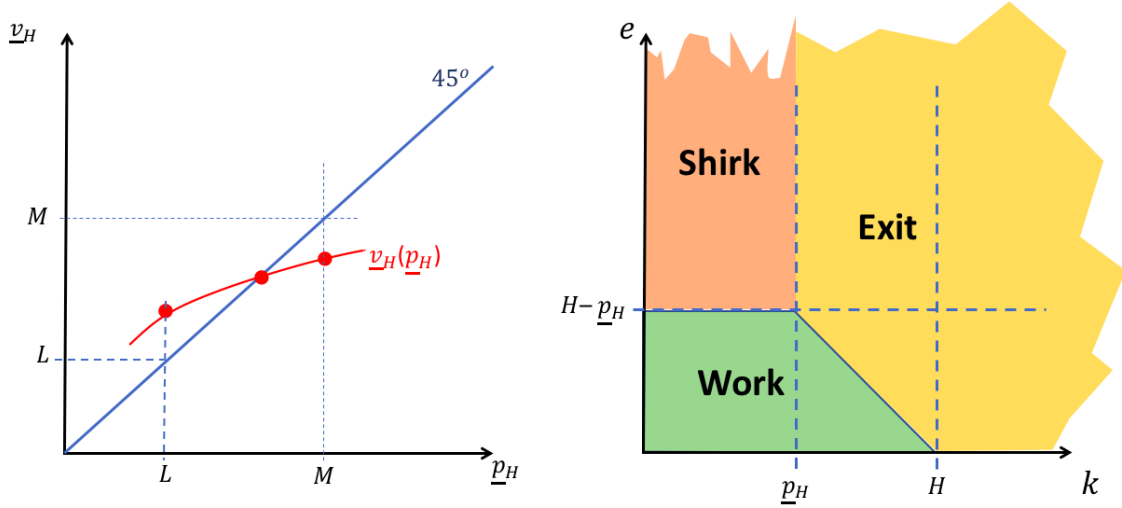
When  $B = H$ ,  $s$ -type sellers will be badged if and only if they choose to exert effort.

**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 prefer to enter, exert effort and be badged over not being badged. Finally, because of unbounded support of  $F(\cdot)$ , and because  $\bar{p}_H - \underline{p}_H$  is bounded, then not all  $s$ -types who enter will exert effort.  $\square$

To characterize the 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  $\bar{p}_H = H$ , 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  $k + e < H$  then they prefer to enter and work over exit (shirking yields negative profits), while if  $k + e > H$  then they prefer to exit. This observation helps characterize equilibrium as follows:

Figure 1: Equilibrium when  $B = H$



**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  $\ell$ - 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  $v_H > \underline{p}_H = L$ , so this cannot be an equilibrium. Now define  $v_H(\underline{p}_H)$  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  $v_H(\underline{p}_H)$  for any  $M > \underline{p}_H > L$  as follows:

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

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

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:<sup>8</sup>

1. **prices:**  $\bar{p}_H = \bar{v}_H = H$ , and  $\underline{p}_H = \underline{v}_H \in (L, M)$ ,
2. **entry:**  $\mu_{\ell H} = G(\underline{v}_H)\mu_{\ell}$ ,  $\mu_{sH} = G(\underline{v}_H)\mu_s + \int_{\underline{p}_H}^H \int_{H-\underline{p}_H}^{H-k} dG(x)dF(y)dx dy$ , 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  $B = M$ .

### 3.3 Comparative Statics

We now compare what happens when the badging requirement becomes more stringent, and changes from  $B = M$  to  $B = H$ .

Comparing equilibrium entry conditions in the two cases we can see that, according to our model, aggregate entry and entrants quality may either increase or decrease with a more stringent budget, depending on the exact shape of the types distribution functions  $G$  and  $F$ .

The following four corollaries follow immediately from comparing prices across the two equilibria and leads to additional empirical predictions.

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

Hence,  $s$ -types who lose their badge are hurt by facing a lower price, and 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.** *a)  $\underline{p}_H > \underline{p}_M$  and  $\bar{p}_H > \bar{p}_M$ ; b) entry increases for quality levels at the tails of the quality distribution and decreases for quality levels in the middle of it.*

This implies that both for  $\ell$ - 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.

---

<sup>8</sup>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.

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

This follows immediately from the fact that all  $s$ -types choose to shirk when  $B = M$  while they must work when  $B = H$ .

**Corollary 4.** *Let market  $A$  have measure  $\mu_s^A$  and let market  $B$  have measure  $\mu_s^B > \mu_s^A$ , fixing the other measures of  $\ell$ - 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, then 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.

This last corollary is critical in generating the main comparative static that guides our empirical analysis. Naturally, when there are more  $s$ -types in a market, then more sellers will necessarily lose their badge after an increase in stringency. Hence, if a policy change occurs, then one can *infer* that market  $B$  had a higher measure of  $s$ -types than market  $A$  if a larger fraction of sellers lose their badge in market  $B$ . Hence, if a policy change is implemented simultaneously in many markets, then corollary 4 implies that in those markets that lost a higher fraction of badged sellers, the impact on the tails of the distribution of entry will be larger, an insight we take to our data.

Here, we summarize the main empirical predictions of the model: i) Depending on the exact shape of distribution functions  $G$  and  $F$ , the total entry rate and the average quality of entrants may increase or decrease after the policy change. ii) Perceived quality and prices should increase for both badged and unbadged sellers. iii) The quality distribution of entrant sellers exhibits thicker tails; conversely, the quality distribution of incumbent sellers who exit has thinner tails. iv) Incumbents who would lose their badge but instead retain it must increase their quality. iv) Across markets, in a market with more  $s$ -types, there will be a larger impact on the tails of the quality distribution of entrants (and exiting incumbents).

## 4 Data

We use proprietary data from eBay that include detailed characteristics on product attributes, listing features, buyer history, and seller feedback. Our data cover the period from October 2008 to September 2010, which include all listing and transaction data in the year before and the year

after the policy change. An important feature of our data is information on product subcategories cataloged by eBay. There are about 400 subcategories (which we also call markets), such as Lamps and Lighting, Beads and Jewelry Making, Video Game Memorabilia, Digital Cameras, and others. A subcategory is the finest level of eBay’s catalog that includes all listings on the site.

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 a seller’s first listing in different subcategories on eBay. We treat this date as a seller’s entry date into the subcategory. 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.

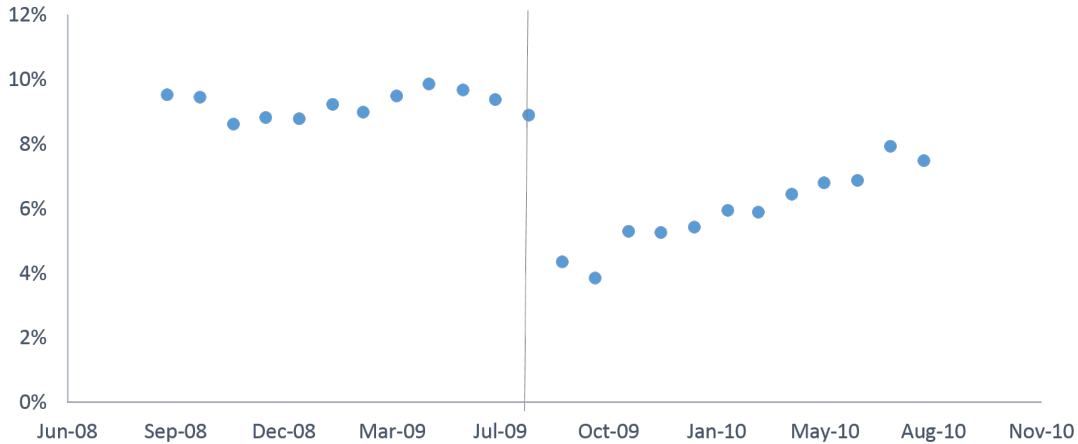
Finally, the use of internal data allows us to construct a measure of quality that is not observed with information that appears publicly. Every seller has a reputation score and percent-positive (PP) on eBay, the latter being the number of positive ratings divided by the total number of ratings. [Nosko and Tadelis \(2015\)](#) demonstrate the extreme skewness of PP, where the mean is 99.3% and the median is 100%, a finding consistent with others who documented biases in reviews ([Zervas et al., 2015](#); [Luca, 2011](#); [Fradkin et al., 2017](#)). [Nosko and Tadelis \(2015\)](#) show that silence is itself a sign of some negativity, and they construct a new measure they call “Effective Percentage Positive” (EPP), which is the number of positive feedback transactions divided by the number of total transactions. [Nosko and Tadelis \(2015\)](#) show 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

The policy change described in section 2 offers a quasi-experiment, and Figure 2 clearly shows that the policy change caused a significant decrease in the share and number of badged sellers. The average share of badged sellers dropped from around 10% during the year before the change to about 4% right after the change, with a gradual re-adjustment taking place in the following year.

We take advantage of the fact that a “one size fits all” policy change was implemented across heterogeneous markets, each having its own distribution of sellers as modeled in Section 3. Our goal is to create treatment and control groups using variations in policy exposure across different markets on eBay. Consider two such markets; after the policy change, one market loses a larger fraction of its badged sellers than the other. Through the lens of the model and Corollary 4,

Figure 2: Share of Badged Sellers



*Notes:* Average monthly share of badged sellers on eBay. The vertical line indicates the policy change after which it was harder for a seller to obtain a badge. All the averages are statistically different from each other at the 1% level.

variation in the intensity of how many sellers lose their badge is an indication of how many *s*-type sellers there are. It follows that a market with a larger drop in the share of badged sellers should exhibit a larger change in outcome variables. Also, we assume that this variation is exogenous to other aspects of a market aside from the distribution of types and test this assumption with different measures of policy exposure in the online appendix as well as using a placebo test.<sup>9</sup>

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.<sup>10</sup>

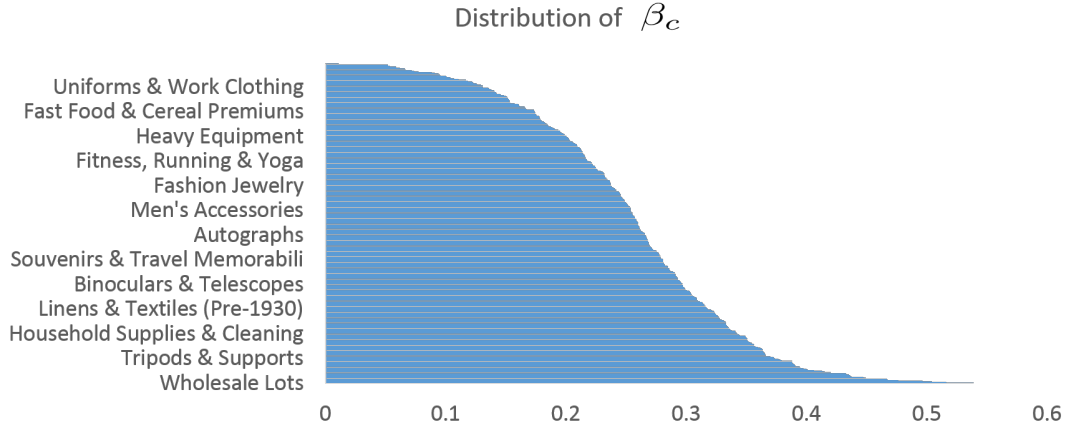
The horizontal bars in Figure 3 are the ex-ante measure of policy exposure, which is the simulated percentage drop in the share of badged sellers across markets. The figure shows that the decrease in the share of badged sellers after the policy change varies dramatically across markets, from under 10% to as much as 50% across markets.

Our main identification strategy exploits the variability in policy exposure in different markets

<sup>9</sup>A similar approach is used in Mian and Sufi (2012).

<sup>10</sup>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 the 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 may endogenously depend on changes in entry due to the policy change.

Figure 3: Policy Exposure in Different Subcategories



*Notes:* Policy exposure is the percentage drop in badged sellers caused by the policy 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 inter-temporal changes in the number and quality of entrants in markets that are more affected by the policy change against the inter-temporal changes of these two measures in markets that are less affected over the same time period. This DiD approach is continuous in the sense that the “treatments” (i.e., policy impacts on the share of badged sellers across markets) take continuous values between 0 and 1. Specifically, the DiD specification is given as

$$Y_{ct} = \gamma \hat{\beta}_c Policy + \mu_c + \xi_t + \epsilon_{ct}, \quad (1)$$

where  $Y_{ct}$ 's are the outcome variables of interest in subcategory  $c$  in month  $t$  (e.g., quality, or entry);  $\hat{\beta}_c$  is the simulated policy impact on the share of badged sellers from our first stage shown in Figure 3;  $Policy$  is a dummy variable that equals to 1 after the policy change;  $\mu_c$  are subcategory fixed effects;  $\xi_t$  are month fixed effects; and  $\epsilon_{ct}$  are error terms.

Our coefficient of interest is  $\gamma$ , 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 positive  $\hat{\gamma}$  means that a larger decrease in the share of badged sellers increases the outcome variable.

As mentioned in the introduction, to ensure that possible endogeneity issues are not driving our results we will then perform two placebo tests as well as a robustness check for the identification.



(see Section 7)

Note that there are two types of entrants: new sellers on eBay (13.3%) and existing sellers entering new markets (86.7%). 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 as entry. In Section 8, we repeat our analyses for the two sets of entrants separately and show that results are similar for the two subgroups.

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  $\hat{\gamma}$  is that serially correlated unobserved errors do not systematically correlate with  $\hat{\beta}_c$  and  $Y_{ct}$  simultaneously. We provide a robustness test of this identification assumption in Table 7 in Section 8.

## 6 Results

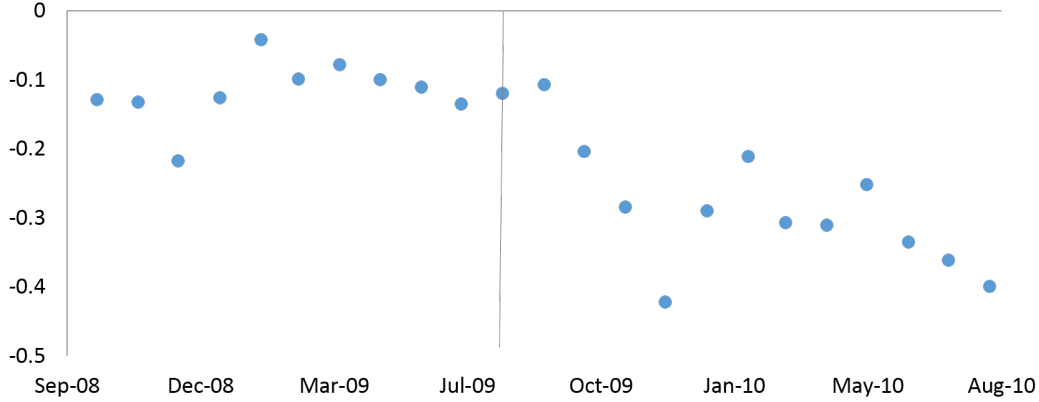
We first estimate the effects of the policy change on the average rates of entry and quality provided by the entrants. We then describe how the quality distribution of the entrants changes. Subsequently, we investigate what can be learned from how incumbents react to the policy change, in comparison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

### 6.1 Effects on Entry Rate and Quality of Entrants

According to our simple model, the total number of entrants can either increase or decrease after the policy depending on the shape of the distributions of types and entry costs. Before conducting the two-stage regression analysis described in the previous section, we 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.

We proceed with some simple descriptive facts by normalizing the entrant ratio and the share of badged sellers for a meaningful comparison. In particular, we compute the percentiles of these two

Figure 4: Market Structure and Entry



*Notes:* Correlation between the percentile of entrant ratio and the percentile of share of badged sellers across subcategories in each month. 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.

measures 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 markets, i.e., markets 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 that the magnitude of this effect is correlated with changes in the share of badged sellers.

Table 1 reports  $\hat{\gamma}$  from regression (1) for entry rate and quality of entrants. Recall that a positive  $\gamma$  means that the increase in the outcome variable is larger in more impacted markets, i.e., a larger drop in the share of badged sellers. In Panel A of Table 1, column 1 shows that the entrant ratio is higher in markets that are more affected, using data from three months before and after the policy change (June 20–September 19 and September 20–December 19 in 2008). 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 (relative to the six months before the policy change), where the estimate is even smaller and is not statistically significant.<sup>11</sup>

In Section 8.2, we show that average price increases more in more impacted markets, which is consistent with our theory and the empirical results on entry.

To understand the distributional impact of the policy change on the number of entrants, Figure

<sup>11</sup>We do not include longer time periods because eBay introduced eBay Buyer Protection in September 2010.

Table 1: Policy Impact on Rate and Quality of Entrants

<i>Panel A. Entrant Ratio</i>			
	(1)	(2)	(3)
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.124*** (0.021)	0.066*** (0.016)	0.010 (0.032)
$R^2$	0.911	0.888	0.685
<i>Panel B. EPP Conditional on Survival in the Second Year</i>			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.064*** (0.019)	0.039*** (0.014)	0.043*** (0.016)
$R^2$	0.771	0.728	0.699

*Notes:* The regressions are at the subcategory-month levels.

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

5a plots two time series, monthly average (normalized) number of entrants and monthly average share of badged sellers, in the markets that are most affected (top quintile of  $\beta_c$ ) and least affected (bottom quintile of  $\beta_c$ ), respectively. Figure 5a shows that in the top quintile, the share of badged sellers decreases from about 35% to less than 15% right after the policy change, whereas in the bottom quintile, 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. This suggests that the policy effect on entry comes mainly from markets that were heavily affected. We show the robustness of these results by looking at top and bottom deciles of  $\beta_c$  in Figure 12.

We now study how the average quality of entrants is affected by the policy change. As for the number of entrants, the model suggests that the average quality may either increase or decrease because of the policy, depending on the distribution.

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 quality 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 8 and show similar patterns.

Positive 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%

Table 2: Two Types of Entry

	Entrant Ratio		Share Badged	
	1 Mo. Before	1 Mo. After	1 Mo. Before	1 Mo. After
New Entrants	0.045	0.044	0%	0%
Existing Entrants	0.295	0.303	11%	4%

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 markets. 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 markets compared to the bottom quintile.

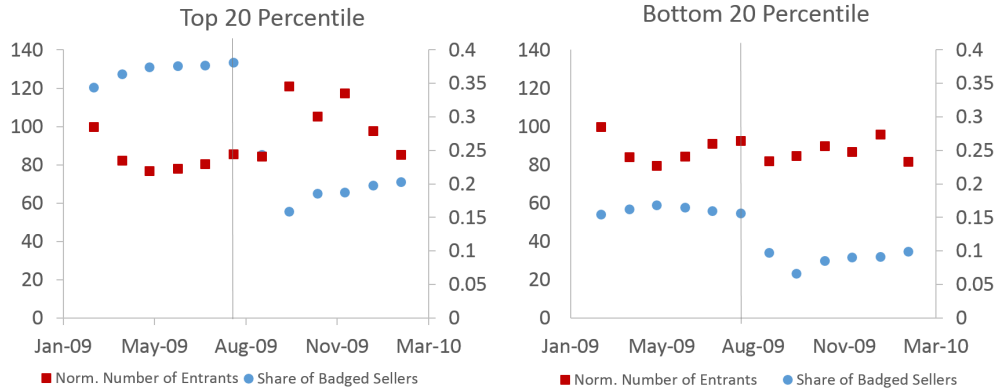
In the online appendix, we also show that there are no significant effects on entrants sales, size and survival rates.

## 6.2 Two Types of Entry

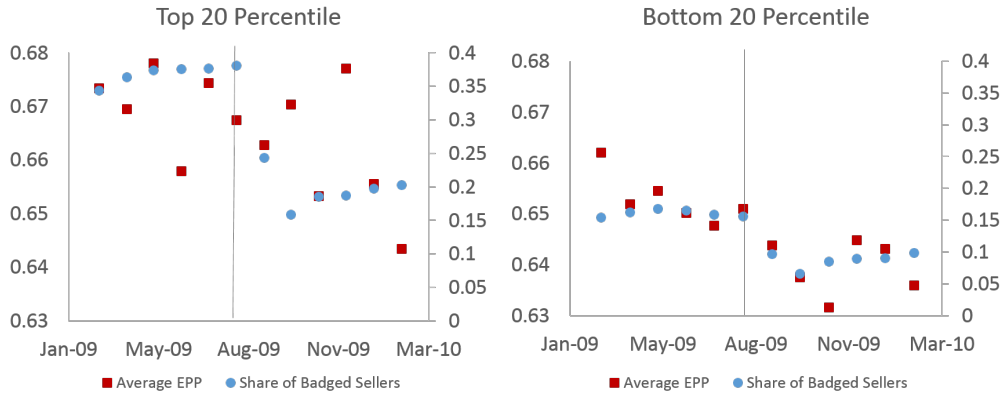
Entrants can be new sellers on the marketplace, as well as existing sellers selling in new markets that they have not operated in before. We find that among entrants into new markets, about 15% are new sellers to eBay and 85% are existing sellers entering a new market. Table 2 shows some summary statistics for these two groups. The first two columns show entrant ratio, the number of entrants divided by the number of incumbents in each subcategory. These numbers do not change a lot when comparing before and after the policy change. The share is around 0.04 for new sellers and 0.3 for the existing sellers. The next two columns, shows the share of entrants of each group that had a badge prior to entering the new categories. By the rules of eBay, no new entrant to the system can be badged upon entry, this shown by the 0% in the first row. On the other hand, when we look at the existing sellers, prior to the policy change 11% of them had a badge and after the policy change 4% of them. This drop echos same drop in share of badged sellers for the average seller as depicted in figure 2.

Through the lens of our theoretical model, these two types of entrants likely differ in their entry cost: the cost of entering eBay is higher than the cost of entering a new market for existing eBay sellers. The former requires sellers to understand the marketplace, its rules and regulations, and

Figure 5: Distributional Policy Impact on Entrants



(a) Distributional Policy Impact on Number of Entrants



(b) Distributional Policy Impact on EPP

*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, respectively. The number of entrants in the six-month period before the policy change is normalized to 100. The number of transactions in the six-month period before the policy change is normalized to 100.

Table 3: 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  $p \leq 0.01$ ; \*\*  $p \leq 0.05$ ; \*  $p \leq 0.10$ .

also to decide which items to sell. On the other hand, the latter only requires that sellers have items to sell in a new category. Differences in the fixed cost of entry will result in differences in the entry decision of the firms as a result of the policy change.

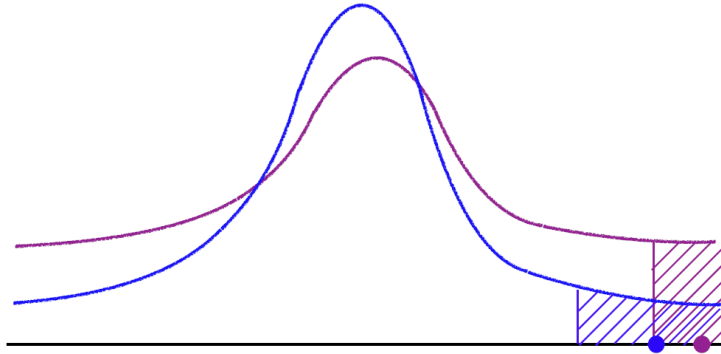
We perform our previous DiD analyses for the two types separately (see Table 3). The relative magnitudes of these estimates are consistent with our theory. If entry costs of starting to sell on eBay are higher than those of entering a new market for an existing seller then new sellers need to have higher 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.

Finally, we regress simulated policy exposure on the share of already badged sellers entering each market, and the estimated coefficient is 0.65, and highly significant. This means that markets that are more affected have more entrants already with certification from other markets, and because this certification is based on their past performance, this can be regarded as a pure selection effect. This suggests that in our data selection is indeed an important determinant of changes in EPP.

### 6.3 Quality Distribution of the Entrant Cohort

Corollaries 1 and 2 of our theoretical framework predict that there should be more entrants of both the highest and lowest quality levels, while entry by middle quality sellers who would have been able to get a badge before the policy but not after will drop. To test these predictions, in each market we partition entrants into deciles based on their EPP score in the first year after their entry.

Figure 6: Example of a Comparison of Two Distributions

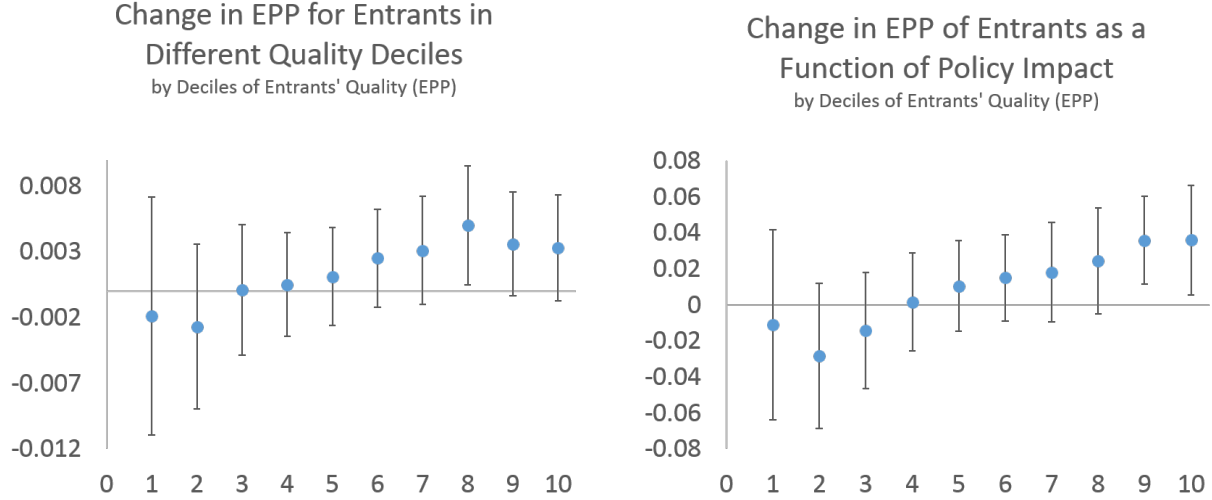


For example, we look at entrants within the top EPP decile based on their transactions in the first year after their entry. We perform two empirical tests, which correspond to the second (*within-market fatter tails*) and fourth (*across-market fatter tails*) empirical predictions of the model. To test for within-market changes in entrants' quality distribution, we rely on an event study approach to estimate the policy effect on EPP for each market (i.e., regressing EPP on a constant, policy dummy, and linear bi-monthly trend), and study the average of these estimates across markets for different deciles. To test for across-market changes in entrants' quality distribution, we perform the DiD specification for different deciles and check if EPPs have increased more in markets more affected by the policy change. In both specifications, when we look at the top decile (entrants with highest quality in the distribution), a positive coefficient will indicate that average entrant quality is higher after the policy, and that average entrant quality is higher after the policy in markets with higher policy exposure, respectively. The intuition for this exercise is shown in Figure 6. When a distribution has a fatter tail, the average of the top decile is higher and to the right.

Figure 7 plots the change in first-year EPP for entrants of different quality deciles. 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 markets with at least 100 entrants (at least 10 entrants in each decile). As a result, for each market we have three observations (six-month equivalent) before the policy change and three observations after it. Additionally, we only consider markets that have entry in all of the six two-month periods and remove markets with a small number of entrants.<sup>12</sup> This leaves us with 228 out of the 400

<sup>12</sup>We have repeated the analysis on all subcategories. We still see the same monotonically increasing estimates as a function of quality deciles, but the results are not significant. This is likely due to the noise induced by having too few entrants in the deciles of some markets.

Figure 7: Change in EPP for Entrants in Different Quality Deciles



Notes: The left figure shows average within-subcategory change in EPP. The right figure shows across-subcategory change in EPP as a function of policy exposure. Bars indicate 95% confidence intervals.

eBay subcategories. The x-axis in Figure 7 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.

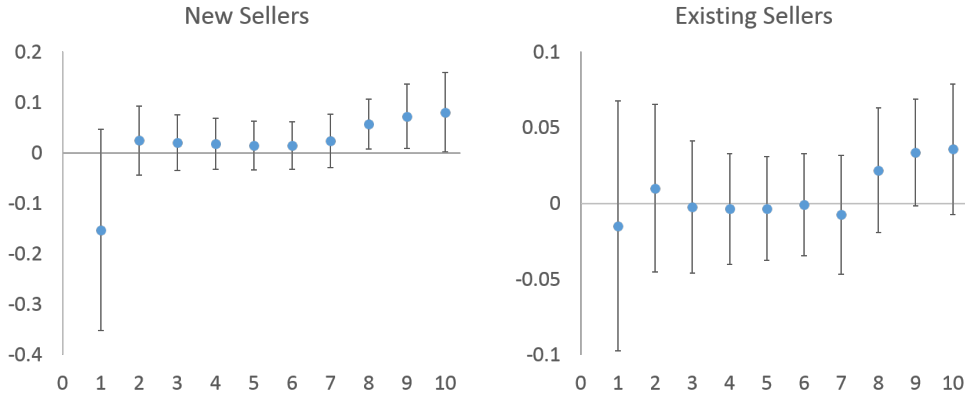
In the left panel we test whether the distribution of entrant quality obtains fatter tails *within* each market. For each quality decile of a market, we estimate how the policy has changed the EPP of entrants in an event-study manner (i.e., regressing EPP on a constant, policy dummy, and linear bi-monthly trend). For each quality decile, the blue dots are calculated by averaging these estimates across markets. Confidence intervals are constructed based on the standard errors of these estimates. In the right panel, we test whether the distribution of entrant quality obtains fatter tails *across* markets. For each quality decile of a market, we perform the DiD estimation across markets, and the blue dots are the estimated  $\gamma$  in specification 1 with their 95% confidence intervals. In both figures, the top-two decile point estimates are positive and statistically significant, as predicted. Though the other estimates are not statistically different from zero, we do observe an overall increasing relationship that is consistent with our prediction that the quality distribution of entrants after the policy change has fatter tails.<sup>13</sup> This in turn implies that sellers in the middle of the quality distribution enter less frequently.<sup>14</sup>

<sup>13</sup>Note that the estimates from the event study approach are an order of magnitude smaller than the ones from the DiD approach. This may be because the DiD approach can better control for common time trends across markets.

<sup>14</sup>We repeated the analyses by dividing entrants into three bins and five bins with qualitatively similar results.



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



Notes: Bars indicate 95% confidence intervals.

Lastly, we have plotted the analogous decile graphs for the two types of entrants separately. Figure 8 shows that we see qualitatively similar findings for both types: we have more entry on the top and bottom deciles of the quality distribution of entrants, although the estimates on the left tail of the quality distribution are not statistically significant.

#### 6.4 Impact on Incumbents

Figure 9 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 to when the policy change took place. Incumbents are defined as sellers who listed at least one item both before and after the change. The EPPs for incumbents are computed using transactions in a given month to capture potential changes in behavior in that month. Entrants in a month are those who have their first listing in that month, and their EPPs are calculated using the transactions in the year after their first listing. The blue series show that there 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 4. 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

Figure 9: Change in EPP of Incumbents and Entrants

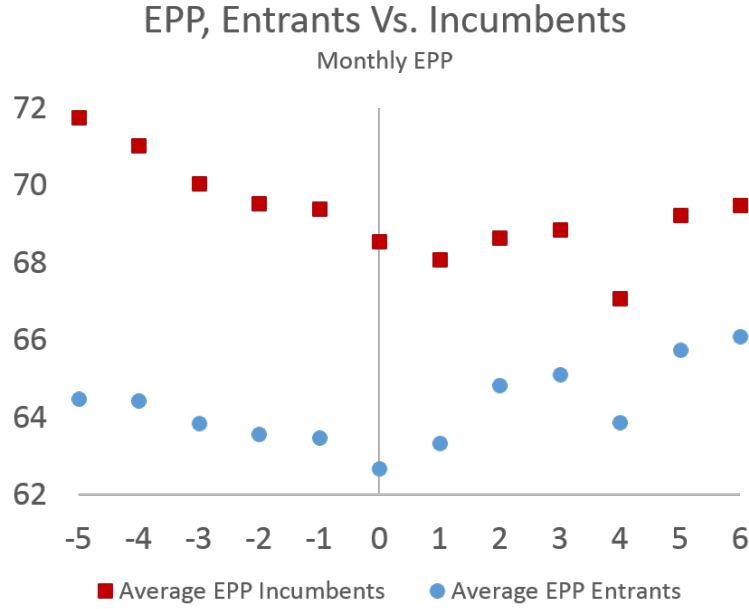


Table 4: Policy Impact on Quality of Incumbents

	(1)	(2)	(3)
<i>Panel A. EPP from Incumbents</i>			
	+/- 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
<i>Panel B. Sellers who Entered <math>n</math> Months before the Policy</i>			
	$n = 3$	$n = 6$	
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  $p \leq 0.01$ ; \*\*  $p \leq 0.05$ ; \*  $p \leq 0.10$ .

positive estimate may suggest that some groups of incumbents increase their quality.

Next, repeat the DiD analyses for sellers who entered not too early before the policy change. These sellers are likely more similar to those that entered right after the policy change because of their proximity in entry date. In Panel B of Table 4, 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, again suggesting that a significant share of the changes in EPP from entrants is likely to come from improved selection.

Corollary 3 stated that incumbents who would lose their badge would have to 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*.<sup>15</sup> A seller was badged before the policy change if she was badged for at least five out of six months before the policy change.<sup>16</sup> 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 5. In Panels A–D, we 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.<sup>17</sup> Using six months before and after the policy change, the only group that experiences a significant increase in EPP in more affected markets 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

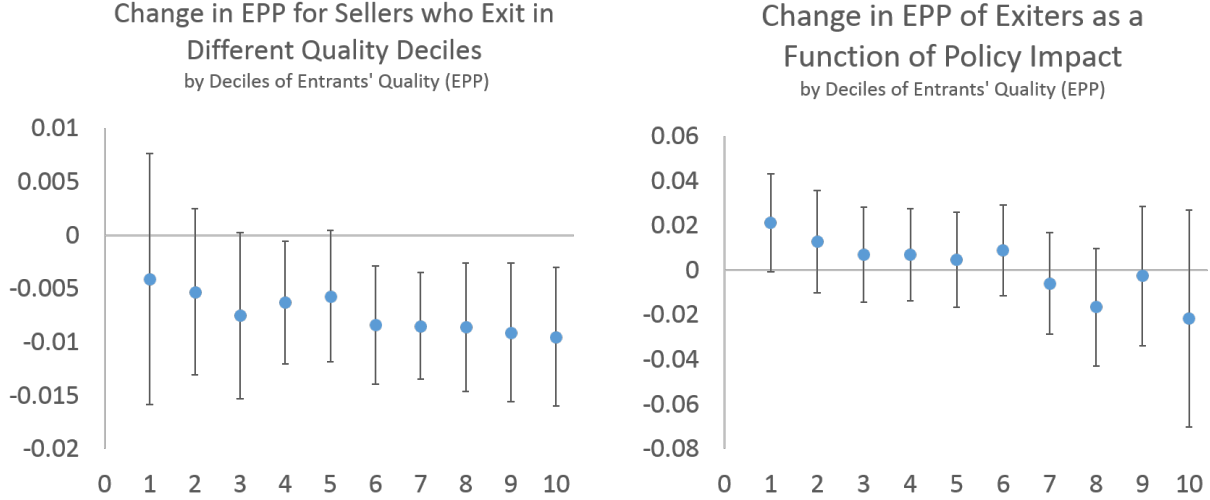
---

<sup>15</sup>The existence of a small group of sellers who were badged only *after* the policy change is due to sellers not being badged instantaneously upon meeting the requirements, but instead being certified once every month.

<sup>16</sup>We considered thresholds for each group of three and four months out of six, yielding qualitatively similar results.

<sup>17</sup>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.

Figure 10: Change in EPP for Sellers who Exit in Different Quality Deciles



Notes: The left figure shows average within-subcategory change in EPP. The right figure shows across-subcategory change in EPP as a function of policy exposure. Bars indicate 95% confidence intervals.

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

The fact that the change in incumbents' behavior is small and is attributed only to a small number of *BN* incumbents once again suggests 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.

Finally, we study changes in the quality distribution of sellers who exit. Figure 10 shows the regression results for each deciles of sellers who exit, with the left figure plotting within-subcategory changes and the right one plotting across-subcategory changes, similar to Figure 7. Decile 1 is for sellers with the lowest quality measured by the EPP. A positive coefficient for decile 1 in the left figure means that the average quality of the lowest decile has increased, and in the right figure means that the average quality of the lowest decile has increased more in more exposed markets, both implying a thinner tail on the left of the distribution in absolute and relative senses. In

Table 5: Policy Impact on Different Incumbent Groups

<i>Panel A. BB Incumbents</i>			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.067 (0.047)	0.048 (0.039)	0.107*** (0.041)
$R^2$	0.661	0.534	0.509
<i>Panel B. BN Incumbents</i>			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-0.018 (0.028)	0.043** (0.020)	0.086*** (0.023)
$R^2$	0.820	0.779	0.753
<i>Panel C. NB Incumbents</i>			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-0.064 (0.059)	0.014 (0.041)	-0.001 (0.044)
$R^2$	0.494	0.473	0.474
<i>Panel D. NN Incumbents</i>			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-0.012 (0.038)	0.007 (0.028)	0.051 (0.031)
$R^2$	0.692	0.648	0.624
<i>Panel E. BN Incumbents who Regain Badge in 3 Months</i>			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.084** (0.041)	0.121*** (0.032)	0.134*** (0.035)
$R^2$	0.705	0.610	0.590
<i>Panel F. BN Incumbents who Remain Unbadged in 3 Months</i>			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-0.044 (0.029)	0.005 (0.022)	0.051** (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  $p \leq 0.01$ ; \*\*  $p \leq 0.05$ ; \*  $p \leq 0.10$ .

this figure, we see that the estimates generally decrease as the quality decile increases, which is the opposite trend of what we have seen in Figure 7. In the right figure, we rely on the DiD specification to control for common time trend across categories, and positive coefficients for bottom deciles and negative coefficients for top deciles imply thinners tail on the left and right of the distribution for the quality of sellers who exit. Although we do not explicitly model exit, this result is the mirror image of the result that the policy change improves incumbents’ outcomes at the tails, thereby reducing their incentive to exit.

## 6.5 Impact on Badge Premium

After the policy change, consumers will see fewer badged sellers in the search result page, possibly changing their valuation of a badge. The price and sales probability of sellers with and without badges may change either because consumers understand the higher quality threshold, or because the 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. (2015) 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 6, we study the change in 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.<sup>18</sup>

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  $BN$  experience an increase in both measures. The magnitudes for both measures in descending order are  $NB$ ,  $BB$ ,  $NN$ , and  $BN$ . Our interpretation

---

<sup>18</sup>Our theoretical model ignores pass-through by assuming that consumers pay their willingness-to-pay.

Table 6: Change in Badge Premium

	(1)	(2)	(3)	(4)	(5)
	Price	Relative Price	Sales Probability	Sales Quantity	Market Share
Policy	-0.035 (0.029)	-0.015*** (0.003)	0.015*** (0.001)	0.010** (0.005)	3.6E-06(21%)* (2.2E-06)
BB*Policy	-0.258 (0.167)	0.010 (0.014)	0.032*** (0.001)	0.035*** (0.012)	3.9E-06(5%*** (5.3E-06)
BN*Policy	-0.517*** (0.074)	-0.052*** (0.006)	-0.014*** (0.001)	-0.031*** (0.004)	-1.1E-05(-16% (2.7E-06)
NB*Policy	0.421 (0.456)	-0.057 (0.041)	0.041*** (0.002)	0.149*** (0.015)	-3.4E-07(-1% (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 small group of sellers in the *NB* group experience an increase in sales because they gain the reputation badge. The sellers in the *BB* group experience an increase in sales because the new badge is more selective 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 sells nothing in a particular week then we fill in zero as her weekly market share. 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 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 after the policy change with similar results reported in the online appendix. .

Summing up the results from Table 6, we conclude that after the policy change the *BN* group is worse off and the other three groups are better off, mostly through increased sales.

## 7 Endogeneity

In this section, we provide evidence that our identification assumption is sensible. An important assumption that we make for the identification step 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 with any other exclusion restriction, we cannot directly test this assumption. Therefore, in this section, we provide here some suggestive evidence that the identification assumption is sensible. To do so we run two placebo tests and also a robustness check for the identification. In addition, we perform an IV estimation to account for cases where the actual change and the simulated change differ.

### 7.1 Placebo Tests on the Exclusion Restriction

Consider the following thought experiment. Suppose there exist serially correlated subcategory-specific confounders that drive our results, and assume that there is some persistency in this confounding effect over time. This 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 perform a placebo test by using the simulated  $\hat{\beta}_c$  and running the second-stage regression on data around September in the *previous* year. Table 7 reports estimated  $\gamma$ 's for entrant ratio, EPP, and total sales for entrants in the previous year, none of which is statistically significant. This implies that the policy change impacted the share of badged sellers in different markets randomly with respect to different entry variables across markets in the previous year. Hence, the policy change generated some exogenous variation in share of badged sellers across markets that are not mere artifacts of heterogeneities across these markets. We also repeat the placebo test in the six months before the policy change, and the estimates are also not statistically significant.



Table 7: Placebo Test on the Exclusion Restriction Assumption

<i>Panel A1: One Year Before the Policy Change;</i>				
$\widehat{\beta}$ Estimated from the Policy Month				
	(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
<i>Panel A2: Six Months Before the Policy Change;</i>				
$\widehat{\beta}$ Estimated from the Policy Month				
	(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
<i>Panel B1: One Year Before the Policy Change;</i>				
$\widehat{\beta}$ Estimated from One Year Before the Policy Month				
	(1)	(2)	(3)	(4)
	Entrant Ratio	EPP	Seller Size	Total Sales
Estimate	3.317	0.024	-2.925	-2852.384
	(2.938)	(0.020)	(3.198)	(25269.507)
$R^2$	0.216	0.765	0.556	0.948
<i>Panel B2: Six Months Before the Policy Change;</i>				
$\widehat{\beta}$ Estimated from Six Months Before the Policy Month				
	(1)	(2)	(3)	(4)
	Entrant Ratio	EPP	Seller Size	Total Sales
Estimate	0.014	0.016	0.029	12912.618
	(0.071)	(0.023)	(2.547)	(22675.753)
$R^2$	0.814	0.729	0.637	0.959

*Notes:* In Panels A1 and A2, we use the  $\widehat{\beta}$  estimated from the year of the policy change, and re-perform the second-stage regression using data around both September in the previous year and March in the policy year. The estimation window used in this table is three months before and after the focal month. In Panel B1 and B2, we use the  $\widehat{\beta}$  estimated from the one year and six months before the policy change, respectively. Other data used in the second-stage is the same as in Panel A's.

We have also performed another placebo test. In particular, we simulate the change in badge requirements at a different date—one year or six months before the actual policy change. These exercises give us two other sets of  $\hat{\beta}$ . We then repeat our regressions around that same placebo date. Estimates in Panels B1 and B2 show that there is no significant changes in outcome variables in these exercises. This result is reassuring because there was no actual change and there should be no impact along the lines that our model predicts.

In principle, there could still exist serially correlated confounders that are not persistent that can contaminate our causal interpretation. However, that the estimates in the placebo test are 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 Controlling for Time-Varying Market Characteristics

Even though we control for fixed effects for different markets, we would have problems if there are time-varying market characteristics that simultaneously correlate with estimated policy exposure and entry. Our placebo tests can detect these time-varying factors only when they are serially correlated. To address un-serially correlated, time-varying market characteristics to the best of our abilities, we rerun our second stage regressions with controlling for many time-varying variables that one might expect impact entry or entrant quality and at the same time might be correlated with our measure of policy exposure. To account for this possibility, we control for the average per category of the following variables: detailed sellers ratings, share of disputes from buyers, percentage of reported defective items, and average EPP for the incumbents. To control for other market characteristics, we control for the following variables in the DiD regression and the result are shown in Table 8. The result are not that different from our baseline result. In fact, the effect of policy seems to have intensified.

## 8 Other Robustness Checks

In this section, we perform several analyses to ensure that our empirical results are robust. First, we show that our result do no change in incumbents' behavior is robust regardless of the time windows used in the definition of EPP. Subsequently, we provide robustness checks on changes in badge premiums for different groups of incumbents by changing the window size of the estimation. Furthermore, we repeat our difference-in-difference analyses with a different first-stage regression.

Table 8: Adding More Controls in DiD Estimation

	Entrant Ratio			EPP		
	(1)	(2)	(3)	(4)	(5)	(6)
	+/- 3 Months	+/- 6 Months	Month 7 to 12	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.135*** (0.022)	0.089*** (0.017)	-0.110*** (0.034)	0.070*** (0.020)	0.035** (0.015)	0.085*** (0.026)
DSR1	0.059 (0.067)	0.019 (0.042)	0.054 (0.057)	0.145 (0.090)	-0.104*** (0.040)	-0.285*** (0.057)
DSR2	-0.193*** (0.068)	0.033 (0.048)	-0.003 (0.087)	-0.141 (0.087)	0.155*** (0.038)	0.097 (0.072)
DSR3	0.233*** (0.052)	-0.107*** (0.030)	0.054 (0.073)	0.147*** (0.054)	-0.054 (0.036)	0.048 (0.057)
DSR4	0.110* (0.057)	0.099*** (0.035)	-0.078 (0.060)	-0.202*** (0.057)	0.063* (0.035)	-0.055 (0.050)
%badged	-0.036** (0.015)	0.000 (0.010)	-0.262*** (0.021)	-0.020 (0.013)	-0.003 (0.009)	-0.030 (0.019)
Price	5E-06 (5E-06)	7E-06*** (2E-06)	-8E-07 (1E-06)	3E-04*** (1E-04)	4E-04*** (6E-05)	-3E-04*** (6E-05)
Quantity	-2E-07 (1E-07)	-3E-08 (4E-08)	-5E-07*** (1E-07)	2E-08 (1E-07)	3E-08 (4E-08)	-4E-08 (5E-08)
# Seller	4E-06*** (8E-07)	3E-06*** (3E-07)	-1E-06 (1E-06)	-1E-06* (7E-07)	-4E-07 (3E-07)	-2E-07 (5E-07)
# Buyer	4E-07 (3E-07)	9E-08 (8E-08)	9E-07*** (3E-07)	4E-08 (2E-07)	-5E-08 (8E-08)	1E-07 (1E-07)
EPP	-0.010 (0.060)	0.027 (0.034)	0.286*** (0.068)	0.112** (0.055)	0.066** (0.031)	0.170*** (0.047)
% Claim	0.490 (0.462)	0.791*** (0.256)	-0.023 (0.452)	-1.666*** (0.445)	-0.338 (0.249)	1.715*** (0.281)
% BBE	0.394 (0.388)	-0.587 (0.249)	0.252 (0.429)	1.647*** (0.433)	0.402 (0.253)	-1.613*** (0.263)
$R^2$	0.918	0.896	0.697	0.778	0.725	0.778

Notes: DSR1-4 are Detailed Seller Ratings for item as described, communication, shipping speed, and shipping charge, on a five-point scale. %badged is number of transaction sold by badged sellers divided by total number of transactions. Price is the average sales price of items in a market. Quantity is the total number of items that are sold in a market. # Seller is the total number of sellers in a market. # Buyer is the total number of buyers in a market. % Claim is the number of disputes filed by buyers divided by total number of transactions. % BBE is the number of bad buyer experience (non-postive feedback, one or two star DSRs, buyer dispute) divided by total number of transactions.

In particular, we use an event-study approach to estimate the actual change in share of badged sellers, instead of using simulation. Additionally, we have also performed an IV estimation to account for cases where the actual changes in share of badged sellers is different from the simulated ones. In the online appendix, we also report results using different window lengths for the first-stage event-study approach, as well as using a normalized rank-preserving measure of  $\beta_c$  in the first stage.

## 8.1 Alternative EPP Measure for Incumbents

We now defined EPP for incumbents based on their transactions in the six months before a particular month. In Figure 13, we plot average EPP for incumbents and entrants using this new definition of EPP and see an increase in the EPP of entrants and no obvious change in incumbents’s EPP.

In Figure 14a, we plot the average *monthly* EPP for incumbents in the four groups based on this new definition of EPP. The solid line is the average monthly EPP provided by incumbents from a particular group in the six months before and after the policy change. 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. 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 because buyers are less likely to leave feedback over time. This implies that the change in the average monthly EPP observed in these two figures maybe due to seasonality.

We created a similar plot for sellers of different quality quartiles measured by EPP in Figure 14b. 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.

## 8.2 Change in Price Across Different Categories

We have established that more entry occurs in markets with a larger drop in the share of badged sellers. Our model argues that more entry is due to a larger price increase (Corollary 4). In this section, we empirically test whether the increase in average price is higher in more exposed markets.

To establish an “apples-to-apples” comparison of prices, we focus on unique Product ID’s—eBay’s finest catalogue that is only defined for homogeneous products, thereby excluding heterogeneous products and construct a dataset at the Product ID–month level. We then perform a DiD estimation

Table 9: Policy Impact on Price in Different Categories

<i>Dependent Variable: Relative Price</i>			
	(1)	(2)	(3)
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.063*** (0.006)	0.094*** (0.005)	0.295*** (0.009)
$R^2$	0.445	0.394	0.514

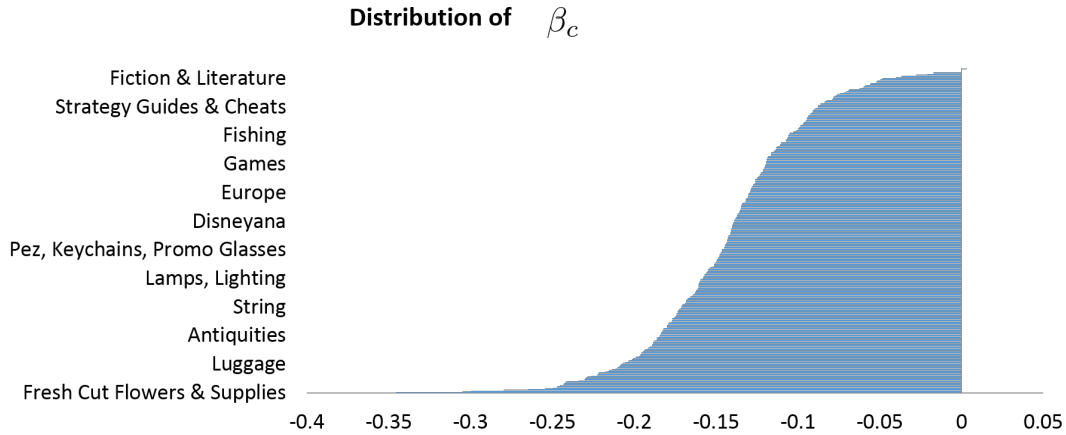
*Notes:* The regressions are at the Product ID-month levels.

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

across product IDs. To facilitate the price comparison, we look at new items only and normalize the prices with respect to the product’s relative price as defined earlier. The positive coefficients in Table 9 mean that the average price is higher in more exposed categories, which is consistent with our theory and our empirical findings.

### 8.3 Event-Study Approach as First Stage

Figure 11: Heterogeneous Impact of Policy Change Based on Event-Study Approach



*Notes:* The estimates are based on data from six months before and six months after the policy change. There are about 400 subcategories, and the labels on the left are just some examples.

In this section, we repeat the difference-in-difference analyses using a different first-stage estimation method. Instead of simulating the change in the share of badged sellers across different subcategories, we estimate the change in the share of badged sellers in different subcategories using the following event study approach:

$$Share\_Badged_{ct} = \beta_c Policy + \eta_c + \alpha_c t + \epsilon_{ct}, \quad (2)$$

Table 10: Event-Study Approach as First Stage

<i>Panel A. Entrant Ratio</i>			
	(1)	(2)	(3)
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.299***	0.204***	0.047
	(0.041)	(0.027)	(0.051)
$R^2$	0.913	0.889	0.691
<i>Panel B. EPP Conditional on Survival in the Second Year</i>			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.102***	0.066***	0.062**
	(0.034)	(0.023)	(0.026)
$R^2$	0.758	0.717	0.690

*Notes:* The regressions are at the subcategory-month levels.

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

where  $Share\_Badged_{ct}$  is the share of badged sellers in subcategory  $c$  in month  $t$ ;  $Policy$  is a dummy variable which equals 1 after the policy change;  $\eta_c$  are subcategory fixed effects;  $\alpha_c$  is a subcategory-specific linear time trend; and  $\epsilon_{ct}$  are error terms. In the appendix, we report full results for the case where we use data from six months before and six months after the policy change to estimate the first stage policy exposure. However, we have also replicated our results on the average change in entry rate and entry quality using 1) estimates of the first stage three months before and three months after the policy change, 2) estimates of the first stage one month before and after the policy change, 3) the number of badged sellers as the dependent variable in the first stage estimation, and 4) the percentage drop in average share of badged seller in different markets directly computed using data from one week before and one week after the policy change, and 5) the number of entrants as the dependent variable in the second stage estimation (instead of using entrant ratio).

The first stage estimates of changes in the share of badged sellers are reported in Figure 11. The correlation between these estimates and those in Figure 3 in the paper (using the simulation approach) is 0.863, and therefore the two measures mostly agree with each other on the estimated policy exposure. Similar as before, we see that the decrease in the share of badged sellers after the policy change is very different across different markets and for some subcategories.

Table 10 reports the difference-in-difference estimation on average changes in our variables of interest, which is analogous to Table 1 in the paper. We see consistent results that average entry rate and EPP increases in markets with higher policy exposure.

Table 11: Instrumental Variable Estimation

Panel A. First-Stage Estimation			
Dependent Var: Actual Change in $\beta$			
	3 Months	6 Months	Month 7 to 12
Intercept	0.017 (0.002)	0.008*** (0.001)	2.4E-4 (0.001)
Simulated Change	0.346*** (0.007)	0.496*** (0.004)	0.531*** (0.005)
$R^2$	0.490	0.745	0.724
F-Stat	2294	13929	12919
Panel B. Entrant Ratio			
	3 Months	6 Months	Month 7 to 12
Estimate	0.295*** (0.042)	0.215*** (0.027)	0.044 (0.028)
$R^2$	0.912	0.890	0.685

*Notes:* We use simulated change in share of badged sellers in a category as the instrument for the actual change in share of badged seller, which is estimated from the event-study approach.

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

## 8.4 Instrumental Variable Estimation

In this section, we perform an instrumental variable estimation, which is a combination of the simulation approach in the main analyses and the event-study approach. In particular, we use the actual change in share of badged seller (based on the event study approach) in the second stage as the policy exposure, but use the simulated change in share of badged sellers as the instrument for it. This is to account for cases where the actual change and the simulated change in share of badged sellers are different. If markets differ in other dimensions like how easy it is for sellers who lose their badges to quickly regain them, we would expect that impacts on things like entry would be smaller in a market where the predicted reduction in badged sellers did not actually occur. The estimation results are reported in Table 11. In Panel A, we report regression results for the first stage in the IV estimation. We see that actual changes and simulated changes in share of badged sellers are highly correlated, which is expected. Another observation is that F-statistics is very large in the first stage, suggesting that the IV is not weak. Moving to the second stage, estimated changes in entrant ratio and EPP in Panels B and C are very similar to the estimates in Table 10. Lastly, the policy effect on average entrants' size and their size in the following year are reported

in Panels D and E, and they are also similar to the corresponding estimates in Table 10.

## 9 Conclusion

A policy change on eBay made the criteria to obtain a quality-signaling badge more stringent. We exploited this change to offer insights into the effects of a more stringent certification policy on entry, which is generally challenging to study. We exploit the differential impact of the policy change on different markets to document that markets that experience a higher reduction in the share of badged sellers following the policy change exhibit larger entry rates after the change. This effect is temporary, it fades away after seven-twelve months. We then find a significant increase in the overall quality provided by entrants in the more affected markets, which contrary to that of entry rates, persists even from the seventh to the twelfth month after the policy change. Consistent with our theoretical framework, we also find that the distribution of quality provided by entrants has fatter tails after the policy change ; conversely, exit occurs more among sellers with average level of quality. Instead, we find little change in the quality provided by incumbents, except for a small group that were at the margin of losing their badge. This last finding on incumbent and the one that a significantly higher fraction of already badged sellers enters categories more affected by the policy, consistently suggest that a significant part of the observed changes in the quality provided by entrants is likely to be linked to selection. Overall, these findings 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 entrants, and hence for how markets evolve in the long run. Our results shed light on the design of reputation and certification mechanisms in markets with asymmetric information as these impact the number and quality of entrants. The ability to encourage the entry of high-quality sellers is not only important for customer satisfaction from platforms, but could also be important for broader impacts such as innovation.



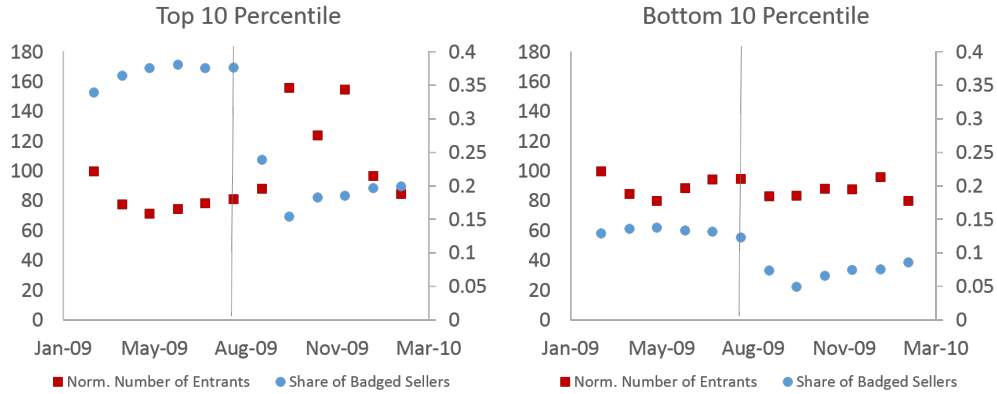
## References

- Avery, Christopher, Paul Resnick, and Richard Zeckhauser**, “The market for evaluations,” *American Economic Review*, 1999, *89* (3), 564–584.
- Bajari, Patrick and Ali Hortacsu**, “Economic insights from internet auctions,” *Journal of Economic Literature*, 2004, *42* (2), 457–486.
- , **Christina Dalton, Han Hong, and Ahmed Khwaja**, “Moral hazard, adverse selection, and health expenditures: A semiparametric analysis,” *The RAND Journal of Economics*, 2014, *45* (4), 747–763.
- Butler, Jeffrey V, Enrica Carbone, Pierluigi Conzo, and Giancarlo Spagnolo**, “Reputation and Entry in Procurement,” *Working Paper*, 2013.
- Cabral, Luis**, “Reputation on the Internet,” *The Oxford handbook of the digital economy*, 2012, pp. 343–354.
- **and Ali Hortacsu**, “The dynamics of seller reputation: Evidence from ebay,” *The Journal of Industrial Economics*, 2010, *58* (1), 54–78.
- Diamond, Douglas W.**, “Reputation Acquisition in Debt Markets,” *Journal of Political Economy*, 1989, *97* (4), 828–862.
- Dranove, David and Ginger Zhe Jin**, “Quality disclosure and certification: Theory and practice,” *Journal of Economic Literature*, 2010, *48* (4), 935–963.
- Einav, Liran, Amy Finkelstein, Stephen P Ryan, Paul Schrimpf, and Mark R Cullen**, “Selection on moral hazard in health insurance,” *American Economic Review*, 2013, *103* (1), 178–219.
- , **Theresa Kuchler, Jonathan Levin, and Neel Sundaresan**, “Assessing sale strategies in online markets using matched listings,” *American Economic Journal: Microeconomics*, 2015, *7* (2), 215–47.
- Elfenbein, Daniel W, Raymond Fisman, and Brian McManus**, “Market structure, reputation, and the value of quality certification,” *American Economic Journal: Microeconomics*, 2015, *7* (4), 83–108.

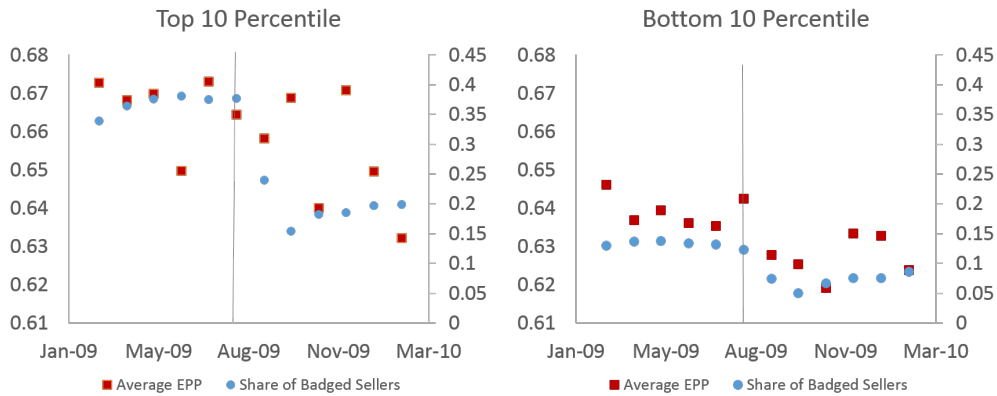
- Fan, Ying, Jiandong Ju, and Mo Xiao**, “Reputation premium and reputation management: Evidence from the largest e-commerce platform in China,” *International Journal of Industrial Organization*, 2016, *46*, 63–76.
- Fradkin, Andrey, Elena Grewal, and David Holtz**, “The determinants of online review informativeness: Evidence from field experiments on Airbnb,” Technical Report, Working Paper 2017.
- (GAO-12-102R), Government Accountability Office**, “Prior Experience and Past Performance as Evaluation Criteria in the Award of Federal Construction Contracts.”
- Greenstone, Michael, Paul Oyer, and Annette Vissing-Jorgensen**, “Mandated disclosure, stock returns, and the 1964 Securities Acts amendments,” *The Quarterly Journal of Economics*, 2006, *121* (2), 399–460.
- Grossman, Gene M and Henrik Horn**, “Infant-industry protection reconsidered: the case of informational barriers to entry,” *Quarterly Journal of Economics*, 1988, *103* (4), 767–787.
- Grossman, Sanford J**, “The informational role of warranties and private disclosure about product quality,” *The Journal of Law and Economics*, 1981, *24* (3), 461–483.
- Hopenhayn, Hugo and Maryam Saeedi**, “Reputation Signals and Market Outcomes,” Working Paper 2018.
- Hui, Xiang, Maryam Saeedi, and Neel Sundaresan**, “Adverse Selection or Moral Hazard: An Empirical Study,” *Journal of Industrial Economics*, 2018.
- , – , **Zeqian Shen, and Neel Sundaresan**, “Reputation and regulations: evidence from eBay,” *Management Science*, 2016, *62* (12), 3604–3616.
- Jullien, Bruno and In-Uck Park**, “New, like new, or very good? Reputation and credibility,” *The Review of Economic Studies*, 2014, *81* (4), 1543–1574.
- Klein, Benjamin and Keith B Leffler**, “The role of market forces in assuring contractual performance,” *Journal of Political Economy*, 1981, *89* (4), 615–641.
- Klein, Tobias J, Christian Lambertz, and Konrad O Stahl**, “Market transparency, adverse selection, and moral hazard,” *Journal of Political Economy*, 2016, *124* (6), 1677–1713.

- Leland, Hayne E.**, “Quacks, Lemons and Licensing: a Theory of Minimum Quality Standards,” *Journal of Political Economy*, 1979, *87* (6), 1328–1346.
- Luca, Micheal**, “Reviews, reputation, and revenue: The case of Yelp. com,” *Harvard Business School Working paper 12*, 2011, *16*, 12–016.
- Mian, Atif and Amir Sufi**, “The effects of fiscal stimulus: Evidence from the 2009 cash for clunkers program,” *Quarterly journal of economics*, 2012, *127* (3), 1107–1142.
- Nosko, Chris and Steven Tadelis**, “The limits of reputation in platform markets: An empirical analysis and field experiment,” Technical Report, National Bureau of Economic Research 2015.
- Shapiro, Carl**, “Premiums for high quality products as returns to reputations,” *Quarterly journal of economics*, 1983, *98* (4), 659–679.
- Stahl, Konrad and Roland Strausz**, “Certification and market transparency,” *The Review of Economic Studies*, 2017, *84* (4), 1842–1868.
- Tadelis, Steven**, “Reputation and feedback systems in online platform markets,” *Annual Review of Economics*, 2016, *8*, 321–340.
- Zervas, Georgios, Davide Proserpio, and John Byers**, “A first look at online reputation on Airbnb, where every stay is above average,” 2015.

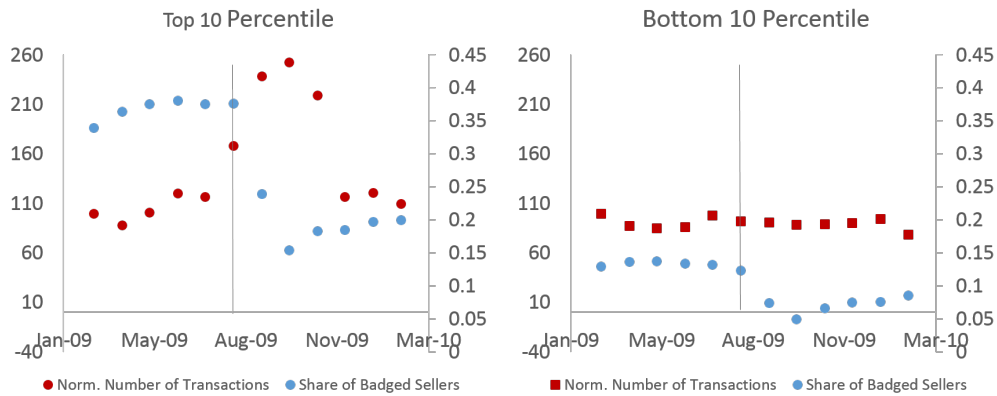
Figure 12: Robustness: Policy Impact on Entrants, Top and Bottom 10 Percentiles



(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 number of entrants in the six months before the policy change are normalized to 100. The number of transactions in the six months before the policy change are normalized to 100.

Figure 13: Robustness: Change in EPP of Incumbents and Entrants

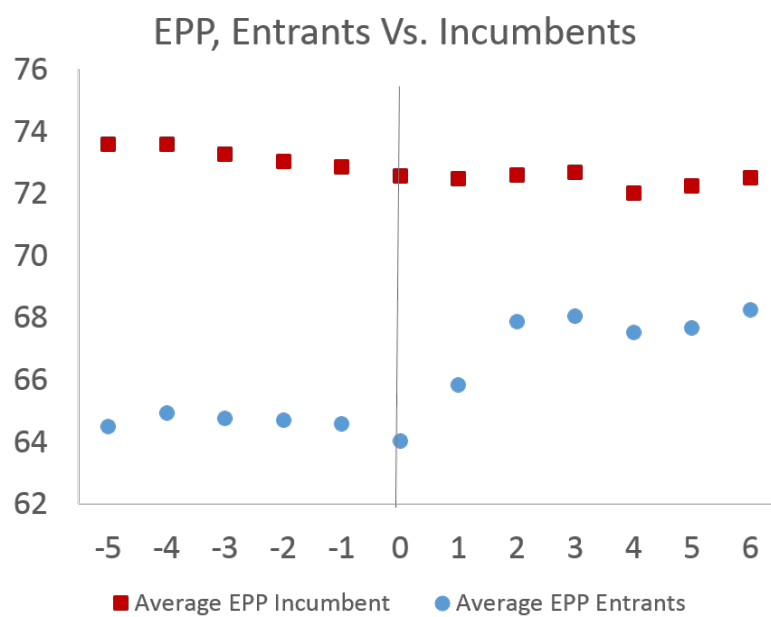
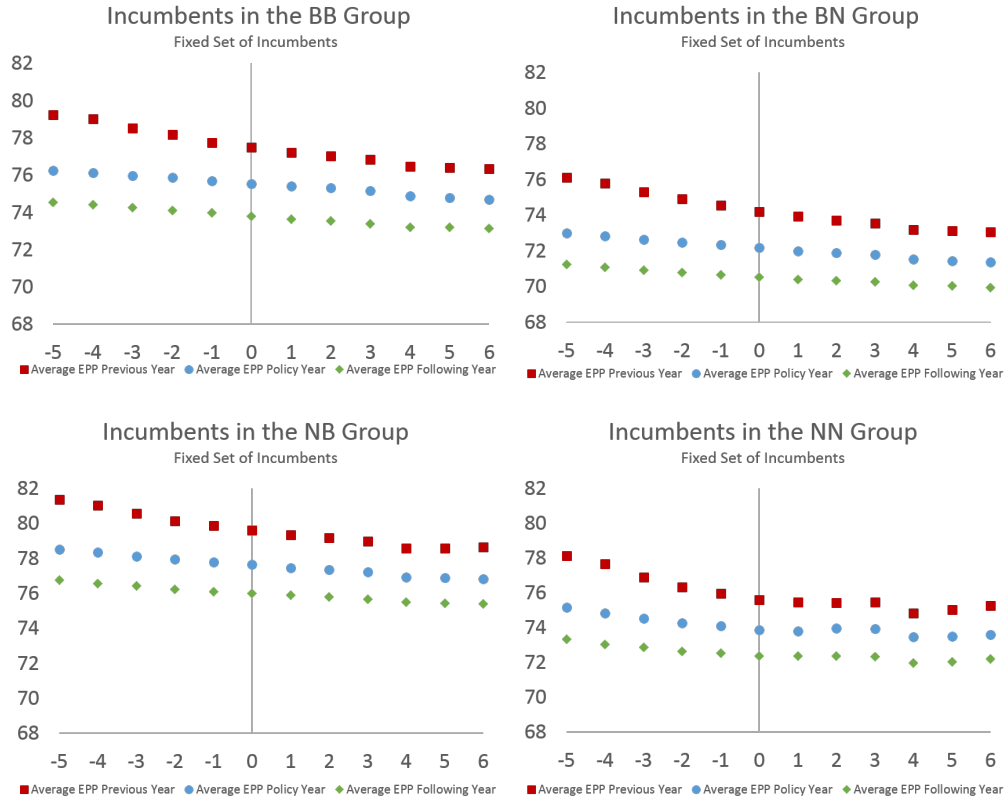
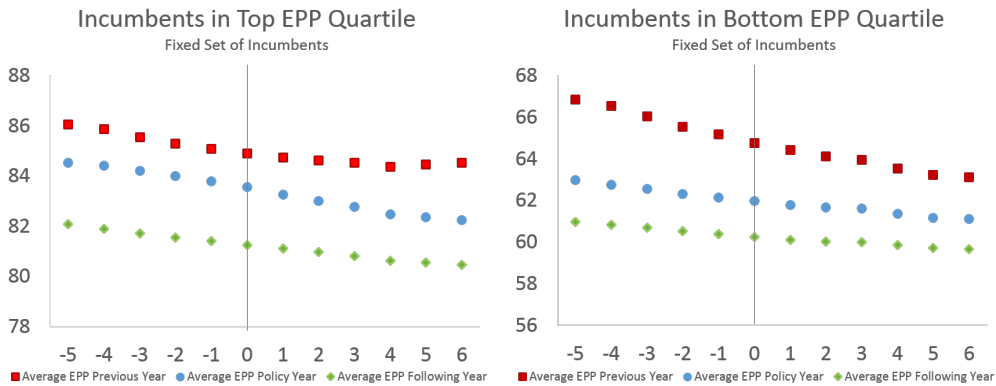


Figure 14: Robustness: Change in EPP of Incumbents



(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.