

# Raising the Bar: Certification Thresholds and Market Outcomes\*

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

August 16, 2019

## Abstract

Certification is a ubiquitous tool that alleviates information asymmetries in markets, yet little is known about the impact of certification-selectivity on market outcomes. Exploiting a policy change on eBay, we study how a more selective certification threshold affects entry, exit, and incumbent behavior. We develop a stylized model that shows how changes in selectivity impact the distribution of quality and prices in important ways. Using rich data from hundreds of online categories on eBay.com, we find support for the model's hypotheses. Our results help inform the design of certification selectivity in electronic and other markets.

*JEL Codes* D47, D82, L15, L86

---

\*We thank David Byrne, Sven Feldmann, Kate Ho, Hugo Hopenhayn, Tobias Klein, 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. An earlier version of this paper was circulated under the title “Certification, Reputation and Entry: An Empirical Analysis”.

<sup>†</sup>hui@wustl.edu

<sup>‡</sup>msaeedi@andrew.cmu.edu

<sup>§</sup>spagnologianca@gmail.com

<sup>¶</sup>stadelis@berkeley.edu

# 1 Introduction

Various 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 intermediary (Leland (1979)). In their extensive survey of quality disclosure and certification, Dranove and Jin (2010) observe that for many important purchases, whether for consumption goods, durable goods, services, healthcare, or schooling, “from cradle to grave, consumers rely on quality disclosure to make important purchases.”

A variety of third-party agencies issue quality certifications, including government agencies (e.g., trade licensing), for-profit rating agencies (e.g., credit ratings), independent NGOs (e.g., green certification), producers’ associations (e.g., sustainable agriculture), and online platforms (e.g., seller quality), to name a few. Like strong brand reputations, certification by a trusted intermediary is often based on past performance and reduces asymmetric information. Furthermore, both strong brand reputations and trusted certification can become barriers to entry for sellers who do not have a certifiable track record (Klein and Leffler (1981), Grossman and Horn (1988)). It seems, therefore, that changes in certification criteria will impact the perceived quality of sellers both with and without certification, and in turn, the resulting market structure mix of incumbents and entrants. How would 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?

In this paper we take a step towards answering these questions, which can help inform regulators and market designers on where to set their certification bar. We begin by developing a parsimonious asymmetric information model of a marketplace in which quality is endogenous and certification affects entry, behavior, and market structure. Using the model’s testable hypotheses, we exploit a policy change that occurred in 2009 when eBay, one of the largest online marketplaces, replaced the “Powerseller” badge awarded to particularly virtuous sellers with the “eTRS” badge that had more stringent requirements, hence “raising the bar” and being more selective.

The model shows that more stringent certification causes the average quality of both badged and unbadged sellers to increase. Sellers who lose their badge are worse than those who remain badged, but are better than those who were not badged previously. As a consequence, when the bar for certification is set higher, entry is encouraged at the tails of the quality distribution, while discouraged in its center. The highest quality sellers benefit from the more selective badge and the lowest quality sellers benefit from being pooled with better non-badged sellers, while entry

becomes less attractive for mid-range quality sellers who can no longer obtain a badge. Hence, our first main testable hypothesis predicts that changing certification stringency will impact the dispersion of quality in the market. A second, more nuanced result shows that markets that are more impacted by the stringency of the badge will display a more dispersed distribution of entrant quality. Finally, the model also shows that only marginal mid-range quality sellers who can increase their quality at a low cost will exert higher effort to profitably obtain the more selective badge.

We take these predictions to the data with an identification strategy that exploits the differential impact of the policy change across 400 separate subcategories (markets) of the eBay marketplace. Through the lens of our model, we assume that the composition of seller quality-types drives the differential impact across markets because the policy change itself was identical in all markets. This leads to heterogeneous effects of the policy on the fraction of badged sellers who lose their badge after the policy change. Indeed, not only do we 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, but we further show that there is substantial heterogeneity of this effect across subcategories.

Using an internally verified measure of quality we find that the distribution of the entrants' quality exhibits "fatter tails" after the policy change, consistent with our theoretical hypothesis. That is, the average quality of entrants is higher in the upper deciles and lower in the bottom deciles of the quality distribution. Furthermore, these results are more pronounced in more affected markets, as our model predicts. We find the opposite effect on exits that exhibit thinner tails, which is the consistent mirror image of our entry results. We also find that more affected markets have significantly more entrants with *pre-existing* certification from other markets, which points at selection as an important driver of changes in quality. Though our model leaves ambiguous whether a more stringent badge will increase entry or average quality, we find that entry increases more in markets where the fraction of badged sellers fell relatively more after the policy change. This effect is significant for the first six months after the policy, after which it fades and becomes insignificant. The average quality of entrants increases significantly after the policy change, and in contrast to the effect on the number of entrants, this increase persists over time.

Because our model predicts that some marginal sellers with intermediate quality levels will change their behavior and provide higher quality, we study the evolution of quality provided by 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.

We then study how prices changed for these four different groups of incumbent sellers—with and without a badge before and after the policy change. The results confirm our model’s predictions: first, sellers who lost their badge experienced a decrease in the relative price that they receive. Second, sellers who remained badged after the change, and those who remain unbadged, experienced higher prices. Third, these changes are more noticeable in more affected markets. To conclude our analysis, we compute a back-of-the-envelope measure of consumer surplus to measure the impact of the policy change. We find that on average, consumer welfare increases by about 2%, yet as expected, is heterogeneous across markets because they exhibit different fundamentals.

An important identifying assumption is that there are no time-varying heterogeneities across subcategories that simultaneously affect changes in the share of badged sellers and in entry. We perform placebo tests and find no impact, consistent with the exclusion restriction of our econometric specification. We perform a series of other robustness tests as well as different specifications of our first stage. These include a flexible event-study approach as well as an instrumental variable approach that combines the estimates of policy exposure from the simulation and event-study approaches. The results are consistently qualitatively similar to those in our main specification.

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>1</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).

Our paper joins a growing literature that uses rich online data to understand how to alleviate asymmetric information in markets. The closest papers to ours are [Elfenbein et al. \(2015\)](#), [Klein](#)

---

<sup>1</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. \(2013\)](#)).

et al. (2016), and Hui et al. (2018), which also used eBay data to study the effects of different information policies on market structure. 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, making it easier for buyers to leave 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 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. 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).<sup>2</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).

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 and offers several robustness tests, while Section 8 concludes the paper.

---

<sup>2</sup>See also Bajari and Hortacsu (2004), Cabral (2012), and Tadelis (2016) for surveys and Avery et al. (1999), Jullien and Park (2014), Stahl and Strausz (2017), and Hopenhayn and Saeedi (2019) for related theoretical papers.

## 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 received a 20% discount on their final value fee (a percent of the transaction price) and had their listings positioned higher on eBay’s “Best Match” search results

---

<sup>3</sup>A Senior Director 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, paying less attention to their performance, which created an incentive for sellers to sell more, sometimes at the cost of 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 more stringent 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 one seller’s listings, signaling the seller’s superior quality to all potential buyers.

Besides the change in the certification policy, two other simultaneous policy changes occurred on eBay.<sup>4</sup> The first introduces easier selling procedures, including faster process for unpaid items, removing buyer feedback if its dispute had been resolved, easier management of buyer messages, and more. The changes apply across all categories, and are controlled for in our DiD approach. The second is a change in the search ranking algorithm, mainly that (i) ranking became based on sales per impression instead of sales; (ii) the title’s relevance was enhanced; and (iii) top rated sellers were promoted in the default search ranking algorithm. The first two changes are controlled for with our DiD approach. For the last change, we include the share of badged sellers in our DiD approach, and still obtain similar results, which are reported in Table ??.<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 two dimensions of cost heterogeneity,  $(k, e)$ , while  $\ell$ - and  $h$ -type sellers only differ across  $k$ .<sup>6</sup>

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

<sup>6</sup>We can alternatively assume that strategic types who do not exert effort will have a baseline quality  $L$  instead of  $M$ . They can then increase their quality to  $M$  by paying the cost  $e$ , or increase their quality to  $H$  by paying a higher

**Demand:** Each of a continuum of buyers demands one unit of a good and is 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  $v_B$  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. Then, if a positive measure of sellers of all types are in the market, then  $\bar{v}_H = H$  and  $M > v_H > L$ , whereas  $H > \bar{v}_M > M$  and  $v_M = L$ .

**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: $B = M$

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

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

The equilibrium when  $B = M$  is therefore characterized as follows:

1. **prices:**  $\bar{p}_M = \bar{v}_M = \frac{\mu_s M + \mu_h H}{\mu_s + \mu_h}$ , and  $p_M = v_M = L$ ,
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.

---

cost  $e' > e$ . Results remain mostly the same except for the prediction that prices increase for unbadged sellers. In this case the price remains the same for unbadged sellers before and after the policy change.



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.

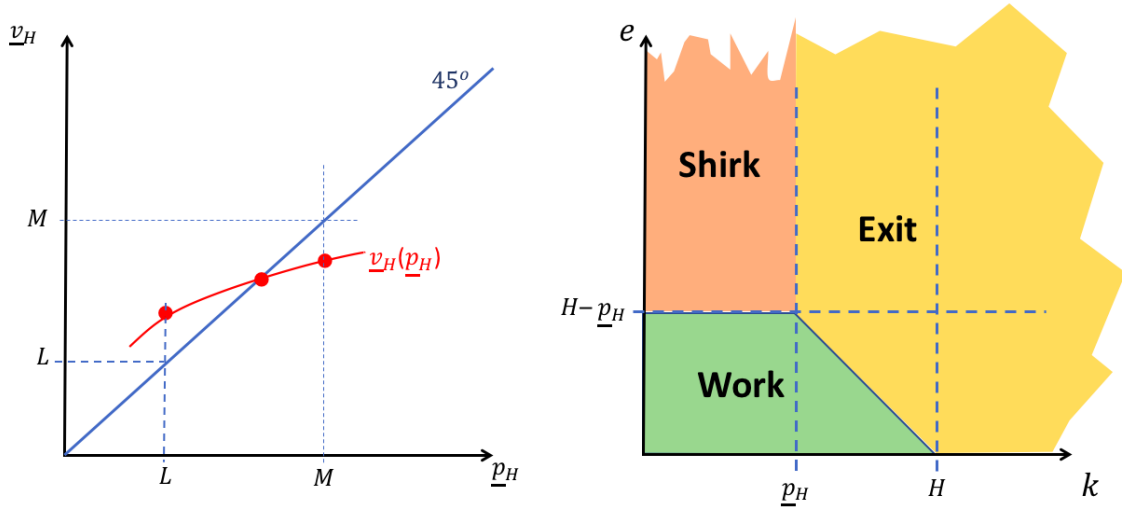
### 3.2 Stringent Badge: $B = H$

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  $G(0) = 0$ , a positive measure of all types will enter the market. This implies that in any equilibrium with  $B = H$ ,  $\bar{v}_H = H$  and  $M > \underline{v}_H > L$ , resulting in  $\bar{p}_H > \underline{p}_H$ . 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 the support of  $F(\cdot)$  is unbounded, and because  $\bar{p}_H - \underline{p}_H$  is bounded, then not all  $s$ -types who enter will exert effort.  $\square$

Figure 1: Equilibrium when  $B = H$



To characterize 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:

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

*Proof.* Market prices determine entry and each  $s$ -type's choice to work. 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  $\underline{v}_H > \underline{p}_H = L$ , so this cannot be an equilibrium. Define  $\underline{v}_H(\underline{p}_H)$  as the unbadged quality that would be obtained following an unbadged price  $\underline{p}_H$  and in which all sellers act optimally. We can explicitly write the function  $\underline{v}_H(\underline{p}_H)$  for any  $M > \underline{p}_H > L$  as follows:

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

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

The left panel of Figure 1 illustrates the logic of Proposition 1. The upshot from the description of equilibria above is that any equilibrium  $B = H$  will satisfy the following:<sup>7</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  $B = M$ .

<sup>7</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.

### 3.3 Comparative Statics

The aggregate entry and entrants' quality may either increase or decrease with a more stringent badge ( $B = H$  instead of  $B = M$ ), depending on the shape of the types distribution functions  $G(\cdot)$  and  $F(\cdot)$ . However, the following five corollaries follow immediately from comparing prices across the two equilibria and lead to testable 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 those with high enough entry and effort costs will not enter after the change.

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

$\bar{p}_H > \bar{p}_M$  follows from the definition of a more stringent badge, and  $\underline{p}_H > \underline{p}_M$  because unbadged sellers now include both qualities  $L$  and  $M$ , which leads to a higher average quality than just  $L$ .

**Corollary 3.** *Entry increases for  $\ell$  and  $h$ -types and decreases for  $s$ -types.*

This Corollary follows directly from Corollaries 1 and 2, which together imply that the distribution of entrants will have “fatter tails” after the more stringent badge is implemented.

**Corollary 4.**  *$s$ -types who retain their badge will increase quality and produce  $H$  instead of  $M$ .*

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

**Corollary 5.** *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.

It is worth noting that instead of only three discrete quality levels and two badge levels we explored a more elaborate model with a continuum of baseline quality-types where each type can increase quality by exerting effort, and the badge can be set at any level of quality in the interior of the type-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 heterogeneity across markets for Corollary 4 would result from heterogenous distributions of types in a similar manner. The model we use is, in our view, the most parsimonious and easy to follow.

We now summarize the main empirical predictions of our model that we take to the data: (i) The quality distribution of entrant sellers exhibits fatter tails, and conversely, the quality distribution of incumbent sellers who exit has thinner tails; (ii) Perceived quality and prices increase for both badged and unbadged sellers; (iii) Incumbents who would lose their badge but instead retain it must increase their quality; and (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 in each subcategory. This 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 quality measure that is not observed 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 (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 \(2015\)](#) construct a measure they call “Effective Percentage Positive” (EPP), which is the number of positive feedback transactions divided by the number of total transactions and show that EPP contains much more information on a seller’s quality than conventional feedback and reputation scores. We follow their approach to compute each seller’s EPP and use it as a measure of quality. 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.

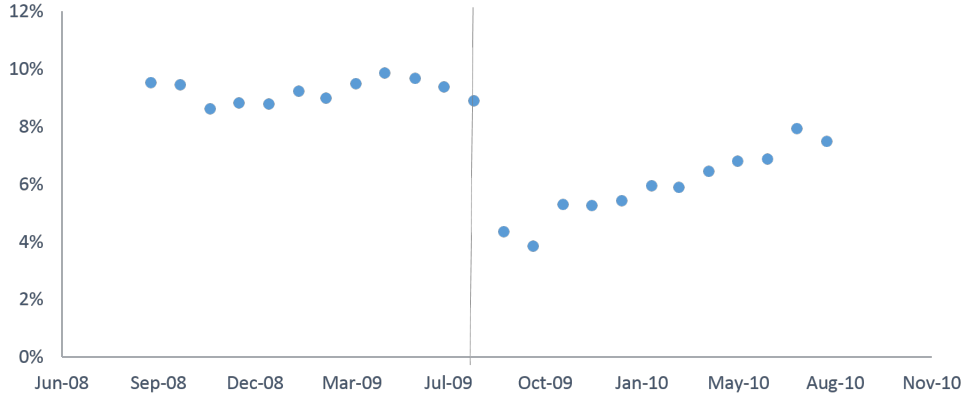
We also considered alternative measures of quality in place of EPP. These include PP, low DSRs, and the number of claims filed against sellers. As we report in the online appendix, all the signs are consistent with our main specification, though some of the regressions are not significant at the 95% confidence level. Also, variations of EPP with different time intervals and without conditioning on survival of sellers yield similar results as reported in the online appendix.

## 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, variation in the intensity of how many sellers lose their badge is an indication of how many  $s$ -type

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.

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. 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>8</sup>

To measure the policy exposure across markets, our first stage is to use the new criteria of a badge 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>9</sup>

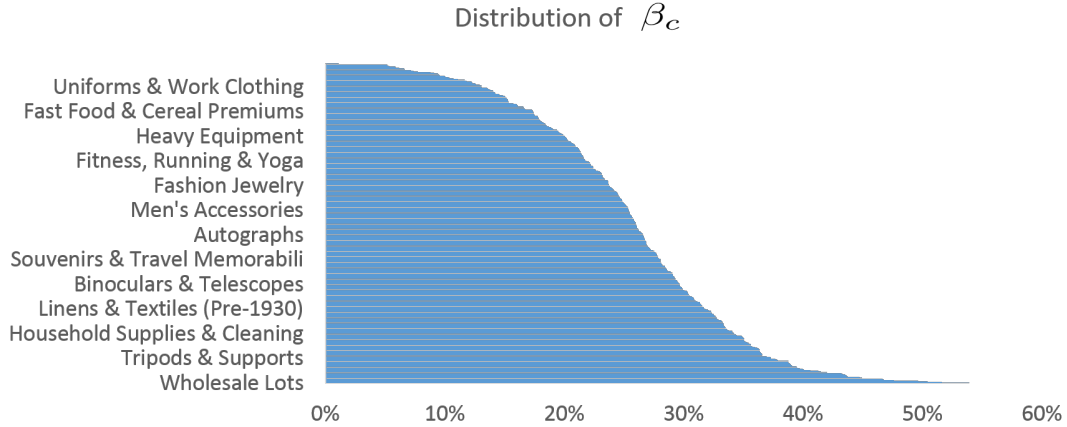
The horizontal bars in Figure 3 are the ex-ante measure of policy exposure, which is the simulated percentage drop in the share of badged sellers across markets. The figure shows that the decrease in the share of badged sellers after the policy change varies dramatically across markets, from under 10%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

Our main identification strategy exploits the variability in policy exposure in different markets induced by the policy change using a continuous difference-in-difference (DiD) approach. In particular, we estimate the policy impact by comparing the changes in the number and quality of entrants

<sup>8</sup>A similar approach is used in [Mian and Sufi \(2012\)](#).

<sup>9</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.

in markets that are *more* affected by the policy change to those 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. The DiD specification is given as,

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

where  $Y_{ct}$ 's are the outcome variables of interest in subcategory  $c$  in month  $t$  (e.g., quality, or entry);  $\widehat{\beta}_c$  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. We cluster standard errors at the subcategory level in the estimation.

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  $\widehat{\gamma}$  means that a larger *decrease* in the share of badged sellers *increases* the outcome variable. Possible endogeneity issues are addressed in Section 7.

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 time-varying 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 6 in Section 7.

Note that there are two kinds of entrants: new sellers on eBay (13.3%) and existing sellers who are “laterally” entering new markets (86.7%). Our theoretical framework implies that these two kinds of entry may behave differently if they differ in their entry costs, which is a reasonable assumption since new sellers need to learn more about how eBay operates. In our main analyses we treat both new sellers on eBay and existing sellers entering new markets as entry. In Section 7.5, we repeat our analyses for the two sets of entrants separately and show that though results are similar for the two subgroups, they are consistent with new sellers having higher entry costs.

## 6 Results

We first present some descriptive statistics of the effects of the policy change on the average rates of entry and quality provided by the entrants, for which our model does not offer sharp predictions, followed by empirical tests of our model’s predictions.

### 6.1 Descriptive Statistics: Entry Rate and Average Quality of Entrants

Our model does not restrict the sign of the change in the total number of entrants, or their average quality, as these depend on the shape of the distributions of types and entry costs. Nonetheless, it is informative to see how these aggregate measures change following the policy change.

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>10</sup>

Positive coefficients in Panel B in Table 1 show that there is an increase in the average quality

---

<sup>10</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***	0.066***	0.010
	(0.021)	(0.016)	(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***	0.039***	0.043***
	(0.019)	(0.014)	(0.016)
$R^2$	0.771	0.728	0.699

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

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

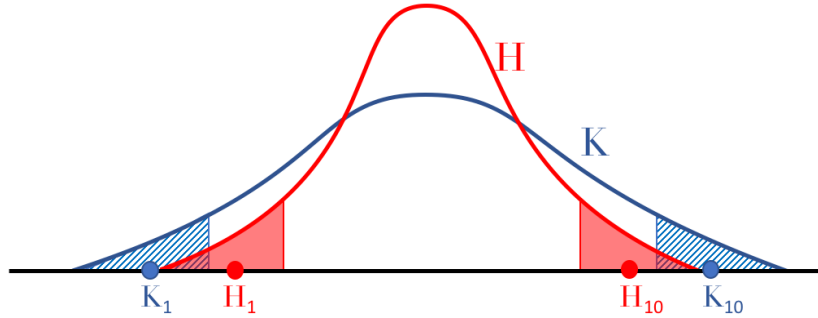
of entrants in the more affected subcategories after the policy change. For a market with a 10% larger drop in share of badged sellers, the policy effect goes from 0.64% to 0.39% as we expand the window length from six (+/- 3) to twelve (+/- 6) months. Column 3 shows that the increase in EPP persists 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.

## 6.2 Quality Distribution of Entrants: Fatter Tails

Two clear predictions of our theoretical model relate to “fatter tails”, which correspond to the first (*within*-market fatter tails) and fourth (*across*-market fatter tails) empirical predictions of the model (see the last paragraph of Section 3). The intuition for testing these is shown in Figure 4. Consider two distributions of entrants' EPP scores in the first year after their entry,  $H$  and  $K$ , the latter having fatter tails. Begin by partitioning entrants into deciles based on their EPP scores. Denote the average quality of the top decile of  $H$  by  $H_{10}$  and of  $K$  by  $K_{10}$ , and similarly denote the average quality of the bottom decile of  $H$  by  $H_1$  and of  $K$  by  $K_1$ . Since  $K$  has fatter tails it follows that  $H_{10} < K_{10}$  and  $H_1 > K_1$ . These differences will be smaller for less extreme deciles and will all but disappear for the middle deciles.

To test for within-market changes in the distribution we rely on an event study approach to estimate the policy effect on EPP for each market, while to test for across-market changes in the distribution we perform our DiD specification for different deciles. Both approaches are explained in more detail below. For both specifications, a positive coefficient for the top deciles indicates that

Figure 4: Example of a Comparison of Two Distributions



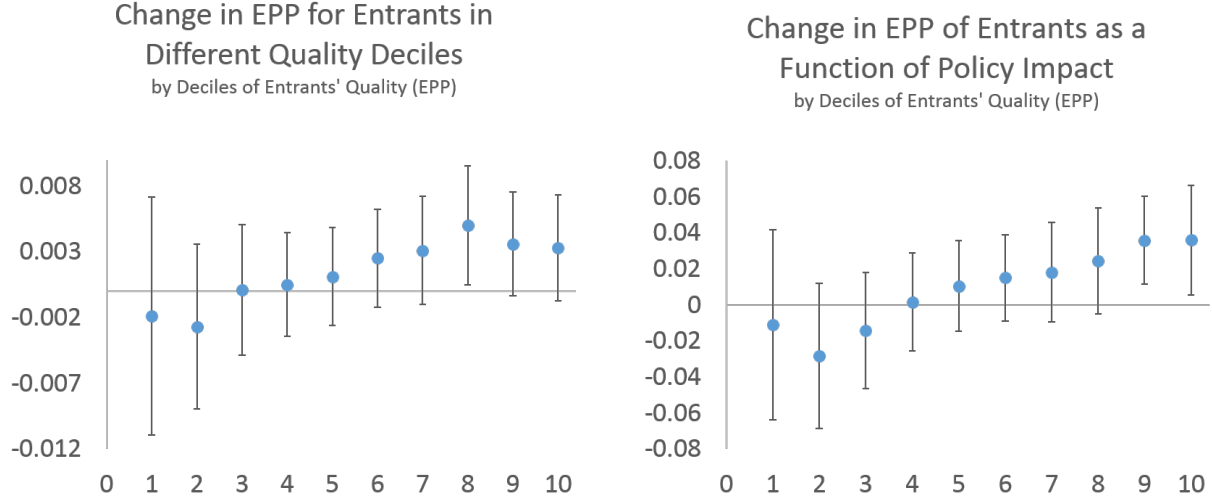
average entrant quality is higher after the policy change, and that average entrant quality is higher after the policy in markets with higher policy exposure, respectively. Similarly, a negative estimate for the bottom deciles will confirm the hypotheses for the bottom tail.

Figure 5 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 and we restrict attention to markets with at least 100 entrants (at least 10 in each decile). Hence, for each market we have three observations (six-month equivalent) before the policy change and three observations after. Additionally, we only consider markets that have entry in all of the six two-month periods, leaving us with 228 out of the 400 eBay subcategories.<sup>11</sup> Figure 5 plots point estimates of the changes in EPP for each decile of the entrant cohorts with 95% confidence intervals, with “10” being the highest decile of EPP and “1” being the lowest decile.

The left panel shows that 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, we plot points 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 plotted points 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

<sup>11</sup>Performing the analysis on all subcategories preserves the monotonically increasing estimates as a function of quality deciles that we find, 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 5: Change in EPP for Entrants in Different Quality Deciles



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

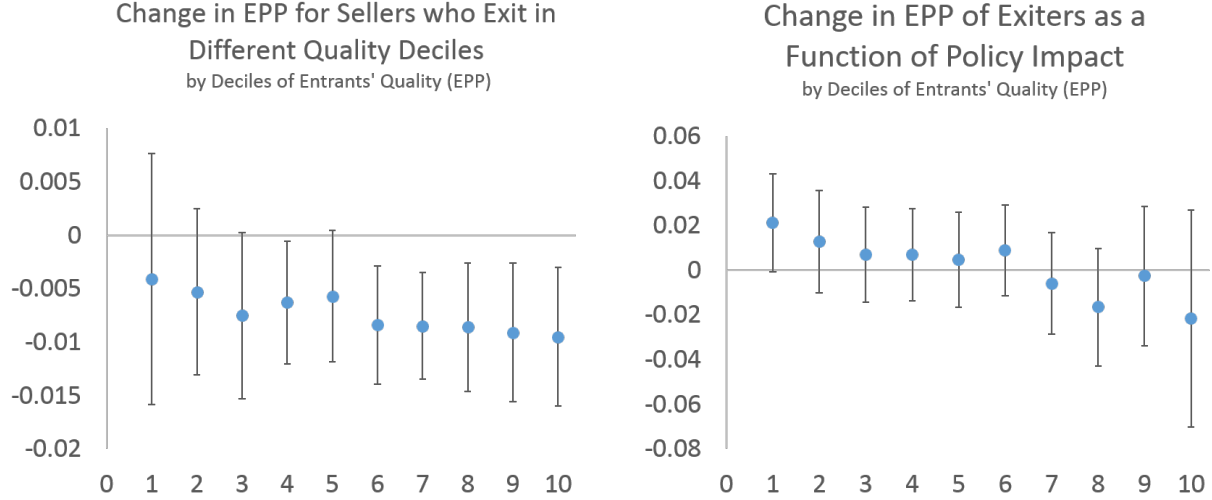
an overall increasing relationship that is consistent with our model’s fatter-tail predictions.<sup>12</sup> This in turn implies that sellers in the middle of the quality distribution enter less frequently.<sup>13</sup>

Finally, we study the natural complement to entry, which are changes in the quality distribution of sellers who exit. Figure 6 shows the regression results for each decile of sellers who exit, with the left panel plotting within-subcategory changes and the right panel plotting across-subcategory changes, similar to Figure 5. 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 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 5. In the right figure, we rely on the DiD specification to control for common time trend across categories. Positive coefficients for bottom deciles and negative coefficients for top deciles imply thinner 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 and offering further evidence consistent with the predictions of our model.

<sup>12</sup>Note that the estimates from the event study approach are an order of magnitude smaller than the ones from the DiD approach, probably because the DiD approach can better control for common time trends across markets.

<sup>13</sup>We repeated the analyses by dividing entrants into three bins and five bins with qualitatively similar results.

Figure 6: 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.

### 6.3 Incumbent Behavior: Some Higher Effort

Our third empirical prediction stated that Incumbents who *would* lose their badge but manage to retain it after the change must have exerted some investment or effort to increase their quality. We define incumbents as sellers who listed at least one item both before and after the change. Incumbents' EPPs are computed using transactions in a *given month* to capture potential changes in behavior and quality of service from month to month.

We start by aggregating all incumbents and run the DiD regression for the quality of incumbents, the results of which are shown in Table 2. In Panel A, we see that although the policy change seems to increase incumbents' EPP in markets with higher exposure to the policy, the changes are not statistically significant at the 10% level. We repeat the DiD analyses for sellers who entered not too early before the policy change as they are likely more similar to those that entered right after the policy change. Panel B of Table 2 shows that there are little changes in behavior for “young” incumbents after the policy change.

To directly test our third empirical prediction, we divide incumbents into four mutually exclusive and collectively exhaustive groups based on their certification status before and after the policy change. One consists of sellers who were badged both before and after the policy change, labeled group *BB*. Another consists of sellers who were badged before but lost their badge after, labeled

Table 2: 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  $p \leq 0.01$ ; \*\*  $p \leq 0.05$ ; \*  $p \leq 0.10$ .

*BN*. We similarly label groups *NB* and *NN*.<sup>14</sup> We consider a seller to be badged before the policy change if she was badged for at least five out of six months before the change.<sup>15</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. The largest group is the *NN* group with over 50% of sellers, while the *NB* group is the smallest at 4%.

We perform the DiD analyses on the four groups of incumbents in Table 3. In Panels A–D, 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>16</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 our model's prediction (last paragraph of Section 3, point (iii)): 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, distinguish between *BN* incumbents based on whether they regain their badge within the three months after the policy change. We see in Panel E that, a *BN* incumbent who regained her badge in the near future increases her quality in the three and six months after

<sup>14</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>15</sup>We considered thresholds for each group of three and four months out of six, yielding qualitatively similar results.

<sup>16</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.

Table 3: 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:* Regressions at the subcategory-month level. Badge status is simulated by applying the new policy requirements to incumbent sellers defined as sellers who list at least one item both before and after the policy change.

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

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 an even finer test that is consistent with our model's prediction, showing that some of the quality improvement is due to more effort exerted by some incumbent sellers. The fact that the

Table 4: Changes in Relative Prices: Event Study

	(1)	(2)
	+/-1 Month	+/-3 Months
Policy	0.005 (0.004)	0.027*** (0.008)
BB*Policy	0.017*** (0.003)	0.027*** 3 (0.002)
BN*Policy	-0.009*** (0.002)	-0.033*** (0.002)
NB*Policy	-0.005 (0.011)	0.059*** (0.008)
Week FE	✓	✓
$R^2$	0.006	0.004

*Notes:* *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.

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

change in incumbents’ behavior 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.

#### 6.4 Prices: Increases and Badge Premiums

The second prediction of our model is that prices will be higher for both badged and unbadged sellers and the fourth prediction establishes that the increase in average prices for badged sellers is higher in more affected markets (last paragraph of Section 3). A challenge in comparing prices on eBay is that products vary wildly because sellers sell many different items that can be new or used, with potentially high variation in the quality of items with the same title.

To establish an “apples-to-apples” comparison of prices, we follow the literature that studies price changes on eBay (e.g., [Elfenbein et al. \(2012\)](#), [Einav et al. \(2015\)](#) and [Hui et al. \(2016\)](#)), by taking advantage of product ID’s in our data to construct an average price for each product that was listed as a new, fixed-price item that was sold. Product ID’s are eBay’s finest-grain catalogue that is only defined for homogeneous products, thereby excluding heterogeneous products to construct a dataset at the Product ID–month level. For each individual item sold we define its “relative price” as the item’s price divided by the average price of the product.

Columns 1 and 2 of table 4 show the changes in the relative prices for different groups of sellers using transactions from one and three months before and after the policy change, where *NN* is the

Table 5: 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***	0.094***	0.295***
	(0.006)	(0.005)	(0.009)
$R^2$	0.445	0.394	0.514

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

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

excluded group. The positive coefficient on *Policy*, which is significant for the  $\pm 3$  month window, shows that overall relative prices increased for unbadged *NN* sellers. Sellers who lose their badge (*BN*) experience a slight decrease in prices while badged sellers (both *BB* and *NB*) experience a larger increase in prices than unbadged sellers. Table 5 shows the DiD estimates across product IDs. The positive coefficients mean that the average prices are higher in more exposed categories, which is consistent with our fourth prediction.

## 6.5 Welfare Impacts: A Back-of-the-Envelope Assessment

Performing a complete welfare analysis is impossible without structurally estimating a variant of our model, forcing us to make assumptions that will surely raise questions as to the validity of a structural approach. Instead, this subsection offers a simple back-of-envelope analysis to estimate the effect of the policy change on consumer surplus. To do this, we focus on items sold in the auction format on eBay, which accounted for half of the listings on eBay at the time.

Instead, we exploit the fact that the auction format on eBay can be approximated by a second-price auction, in which it is known that a bidder’s dominant strategy is to bid their valuation. Hence, we use the winning bidder’s bid as a proxy for the bidder’s willingness-to-pay, and the differences between the winner’s value and the price she pays therefore yields an approximation for the winner’s consumer surplus. In order to control for the type of goods sold before and after the policy change, we control for the items’ product IDs.<sup>17</sup>

Using weakly time dummies to control for seasonality, we find that the average consumer surplus has gone up by about 2%, which indicates that *on average* the policy change has been beneficial for the consumers. Interestingly, however, the impact of the policy change exhibits a wide dispersion

<sup>17</sup>As mentioned earlier, product IDs are eBay’s finest category of items. For example, an iPhone 6, white, 64GB, contracted with AT&T will have its own product ID, and another iPhone of different color or different internal memory will have a different product ID.



across subcategories of products. Recall that our model shows that the underlying distribution of seller types will determine how the policy change will impact entry and exit, and consumer preferences will determine whether this is welfare increasing or decreasing.

Because we cannot predict what characteristics in each subcategory should indicate an increase or a decrease of welfare after the policy change, we use machine-learning tools, particularly LASSO, for guidance. We included many variables such as our first period estimate of change in share of badged sellers, share of badged sellers before the policy change, share of transactions with a negative outcome, measured using different indicators, average price in the category, and the size of different groups of sellers. We find that categories that have a higher share of badged sellers before the policy change observed the highest increase in welfare, as well as categories which had a higher share of negative outcomes.

## 7 Endogeneity and Robustness

This section offers some support for our identification assumptions as well as analyses that show the robustness of our results to changes in specifications. Recall that a critical assumption we make for identification is that there are no time-varying heterogeneities across subcategories that simultaneously affect both changes in share of badged sellers and changes in entry variables. Like with any exclusion restriction, we cannot directly test this assumption. Instead, we present suggestive evidence that the identification assumption is sensible by running two placebo tests as well as a robustness check for the identification. In addition, we perform an IV estimation to account for cases where the actual policy change and the simulated change differ.

We also perform several analyses to ensure that our empirical results are robust to different specifications, including different time windows used in the definition of EPP, repeating our DiD analyses using an event-study approach instead of using a simulation, and an IV estimation to account for cases where the actual changes in the 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.

Table 6: Placebo Test on the Exclusion Restriction Assumption

<i>Panel A1: One Year Before the Policy Change;</i>				
$\hat{\beta}$ 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>				
$\hat{\beta}$ 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>				
$\hat{\beta}$ 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>				
$\hat{\beta}$ 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:* The estimation window used in this table is three months before and after the focal month. Panels A1 and A2 use the  $\hat{\beta}$  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. Panels B1 and B2 use the  $\hat{\beta}$  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.

## 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 6 reports estimated  $\gamma$ 's for entrant ratio, EPP, and total sales for entrants in the previous year, none of which are 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.

In a second placebo test we simulate the change in badge requirements at a different date—one year or six months before the actual policy change—giving 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.

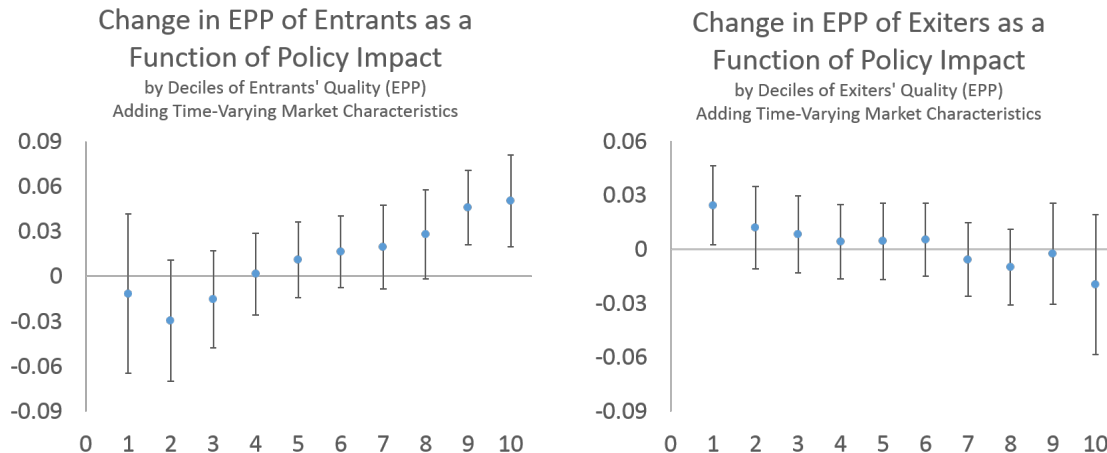
In principle, there could still exist serially correlated confounders that are not persistent that can contaminate our causal interpretation, which we talk about in the following section. However, the fact that the estimates in the placebo test are very noisy is reassuring; for example, the standard 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

Despite controlling for market fixed effects in our DiD specification, identification problems arise 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. One way to mitigate concerns over time-varying, non-serially correlated market characteristics is to rerun our second stage regressions controlling for many time-varying variables that might impact entry and entrant quality and at the same time be correlated with our measure of policy exposure. We can then test whether the estimates are robust to the inclusion of these time-varying market characteristics.

In particular, we control for category-level averages of the following variables: the four detailed sellers ratings, share of badged sellers, average sales price, total number of items sold, total number of sellers, total number of buyers, share of disputes from buyers, percentage of reported defective

Figure 7: Changes in EPP for Entrants and Exitors in Different Quality Deciles, Controlling for Time-Varying Market Characteristics



Notes: The two figures correspond to the right graph in Figure 5 and the right graph in Figure 6, respectively. Both figures across-subcategory change in EPP as a function of policy exposure using the DiD specification. Bars indicate 95% confidence intervals.

Table 7: Policy Impact on Price in Different Categories, Controlling for Time-Varying Market Characteristics

<i>Dependent Variable: Relative Price</i>			
	(1)	(2)	(3)
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.057***	0.068***	0.214***
	(0.007)	(0.006)	(0.010)
$R^2$	0.440	0.399	0.528

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

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

items, and average incumbent EPP.

In Figure 7, we replicate our analyses on changes in entrants and exitors of different quality deciles. The two figures correspond to the right graph in Figure 5 and the right graph in Figure 6, respectively. We see in both graphs that the monotonic increasing and decreasing patterns mostly hold even after controlling for the above-mentioned variables.

Next, we repeat our price analyses in Table 5 by adding more controls. Here the regressions are performed at the Product ID-month level, and therefore the variables are measured at the Product ID level. The estimated policy effect on prices are reported in Table 7. The positive coefficients across different time windows mean that the average prices are higher in more exposed categories, which is consistent with earlier results in Table 5.

Table 8: 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*** (0.041)	0.204*** (0.027)	0.047 (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*** (0.034)	0.066*** (0.023)	0.062** (0.026)
$R^2$	0.758	0.717	0.690

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

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

### 7.3 Event-Study Approach as First Stage

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:

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

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.<sup>18</sup>

The first stage estimates of changes in the share of badged sellers are reported in the online appendix and correlation between these estimates and those obtained by the simulation approach is 0.863, hence leading to very similar results. Indeed, Table 8 reports the DiD estimation on average changes in our variables of interest, analogous to Table 1 in the paper. We see consistent results

<sup>18</sup>We replicated our results on the average change in entry rate and entry quality using (i) estimates of the first stage +/- 3-months around the policy change; (ii) estimates of the first stage +/- 3-months around the policy change; (iii) the number of badged sellers as the dependent variable in the first stage estimation; (iv) 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 (v) the number of entrants as the dependent variable in the second stage estimation (instead of using entrant ratio).

Table 9: 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  $p \leq 0.01$ ; \*\*  $p \leq 0.05$ ; \*  $p \leq 0.10$ .

that average entry rate and EPP increase in markets with higher policy exposure.

#### 7.4 Instrumental Variable Estimation

In this section, we take an instrumental variable approach that combines the simulation approach in the main analyses and the event-study approach. In particular, in the first stage, we regress the actual change in share of badged sellers from the event study on the simulated change in share of badged sellers. In the second stage, we regress outcome variables on the predicted actual change in share of badged sellers.

The logic of the IV approach is as follows. We want to regress entry outcomes on actual change in badged seller, which is our measure for the treatment intensity. But this treatment intensity may be endogenous. Therefore, we need an excluded variable that changes the propensity of receiving the (continuous) treatment, but otherwise uncorrelated with entry. The excluded variable in this approach is the simulated change in badged seller before the policy change. The exclusion restriction assumption is likely to hold: This variable is pre-determined based on the exogenous distribution of seller types and unobservable to market participants, so it should not correlate with entry in the post period, except through its effect on the actual change in share of badged sellers.

Table 10: New versus Lateral Entry - Entrant Ratio and Share Badged

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

The estimation results are reported in Table 9. 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, as expected. Another observation is that the F-statistics in the first stage are very large, 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 8. 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 8.

## 7.5 Lateral versus New Entrants

As mentioned briefly earlier, entrants can either be new sellers on eBay or existing sellers laterally moving into new markets they have not operated in before. We find that among entrants into new markets, about 13% are new sellers to eBay and 87% are existing sellers entering a new market. Table 10 shows some summary statistics for these two groups. The first two columns show the entrant ratio, the number of entrants divided by the number of incumbents in each subcategory, which does not change after the policy change. The entrant ratio 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 is shown by the 0% in the first row. On the other hand, when we look at existing sellers, prior to the policy change 11% of them had a badge and after the policy change only 4% of them did. This drop echoes the 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 kinds 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 also to decide which items to sell. On the other hand, the latter only requires that sellers decide to expand laterally into new subcategory. 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.

Table 11: New versus Lateral Entry - Policy Change Impact

	New Sellers		Lateral Entrant	
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 $p \leq 0.01$ ; ** $p \leq 0.05$ ; * $p \leq 0.10$ .				

We perform our previous DiD analyses for the two kinds separately (see Table 11). The relative magnitudes of these estimates are consistent with our theory. Namely, 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 the simulated policy exposure on the share of already badged sellers entering each market. The estimated coefficient is 0.65, and is highly significant. Hence, markets that are more affected have more entrants that were previously badged. Because certification is based on past performance, this can be regarded as a selection effect, suggesting that for this policy change, selection is indeed an important determinant of increased quality (EPP).<sup>19</sup>

## 8 Conclusion

We develop a parsimonious model of certification in markets with asymmetric information to explore how certification choices impact market outcomes, and in particular, the distribution of quality. We take the model's predictions to data from eBay and exploit the heterogeneous impact of a policy change across different markets, which allows us to identify our model's rich predictions regarding how such a policy change will impact the quality distribution of sellers and prices across markets.

<sup>19</sup>In the online appendix we plot the analogous decile graphs for the two kinds of entrants separately, showing qualitatively similar findings for both kinds of entrants.



The predictions of our theoretical model are borne out in the data. First, we find that the distribution of quality provided by entrants has fatter tails after the policy change, and conversely, exit occurs more among sellers with mid-range levels of quality. Second, we find that most incumbents do not change the quality of their service except for a small group of incumbents who regain their badge by increasing their quality. Furthermore, a significantly higher fraction of already badged sellers enters categories more affected by the policy. This, together with the finding that only a small number of incumbents increase their quality, suggest that a significant part of the observed changes in the quality provided by entrants is linked to selection in entry and exit. Finally, restricting attention to well defined products, we find that aside from the products of sellers who lose their badge, relative prices go up, and this increase is more pronounced in more affected markets.

Overall, our 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. As online marketplaces have become more widespread, from products, to services, to labor markets and more, our results can offer guidance for electronic marketplace designers, where the use of certification mechanisms is commonplace. According to our findings, raising (lowering) the bar of the certifying badge will not only improve (reduce) the quality of certified providers, but will also broaden (contract) the quality distribution of all sellers in the marketplace. The optimal certification bar is dependent on market characteristics and consumer preferences, and as our back-of-the-envelope analysis shows, a well chosen change in the certification threshold can increase consumer surplus, making the marketplace a more attractive place for consumers.

We view our contribution as a first step in providing insights on matters of quality certification that apply more broadly in other markets with asymmetric information where certification is ubiquitous. These include financial markets where credit ratings are used to obtain the “investment grade” badge; service markets in which accreditation certifies a certain level of service quality; final and intermediate goods markets where labelling institutions certify various forms of quality; and public procurement markets where regulatory certification can significantly change the competitive environment and reduce the costs of public services. If anything, the proliferation of online search and the ability of individuals and firms to access larger amounts of information and more possible trading parties should make the use of certification badges even more ubiquitous.

## 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**, “Charity as a Substitute for Reputation: Evidence from an Online Marketplace,” *The Review of Economic Studies*, 2012, *79* (4).

- , – , and – , “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**, “Optimal Quality Ratings and Market Outcomes,” Working Paper 2019.
- 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.