the public feedback and the anonymous feedback of Detailed Seller Ratings (DSR) to show that the improvement in transaction quality is not due to exits from low-quality sellers. Using internal data from eBay, Hui et al. [2017] complement Klein et al. [2016] and further investigate changes in the size of incumbents. They found that although low-quality sellers do not exit after the policy change, their size shrinks dramatically, which accounts for at least 68% of the quality improvement. In comparison with these three papers, our paper explicitly studies the impact of certification on the dynamics of entry and the changes in market structure, as well as the quality provided by incumbents before and after the change.

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

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

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

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

3 Certification and Entry: A Simple Framework

To guide our analysis, we present a simple example based on Hopenhayn and Saeedi [2016]. Assume that a market is perfectly competitive. Firms differ along two dimensions: quality, z, and fixed costs, f. Assume that $z \in \{z_1, z_2, z_3\}, z_1 < z_2 < z_3$, with mass m_1, m_2, m_3, and normalize the total mass of firms to 1. Fixed costs are independently distributed across sellers with cumulative distribution function $G(f)$. Production technology is the same for all firms, and is given by a strictly increasing supply function $q(p)$, the corresponding variable cost $c(q)$, and the variable profit function $\pi(p)$.

A marketplace regulator can produce an observable certification badge that credibly signals if the quality is at least a certain threshold $z^* \in \{z_2, z_3\}$. That is, we consider two cases: when $z^* = z_2$ then the badge signals that the seller is at least of quality z_2, and the more stringent case when $z^* = z_1$ then the badge signals that the seller is of quality z_1. Denote by p_H and p_L the competitive price for firms above and below the threshold, respectively. It follows that the measure (number) of sellers entering from each type will be $n(p) = G(\pi(p))$,where $p \in \{p_L, p_H\}$. Naturally, the number of entrants increases in the price that they receive.

We assume that demand is given by a baseline demand function $P(Q)$ that depends on the total amount of goods of all quality levels, and an additive quality offset $\Delta \bar{z}$ for a good of expected quality $\bar{z}$. Hence, if the total quantity of all goods in the market is Q, then the demand price for a specific good of expected quality $\bar{z}$ is $P(Q) + \Delta \bar{z}$.

An *equilibrium* for threshold $z^* \in \{z_2, z_3\}$ is a pair of prices, p_H and p_L, and quantities, Q_H and Q_L, such that

1. $p_H = P(Q) + \Delta_H(z^*)$,
2. $p_L = P(Q) + \Delta_L(z^*)$,
3. $Q_H = q(p_H) n(p_H) m_H(z^*)$, and
4. $Q_L = q(p_L) n(p_L) m_L(z^*)$,

where $Q = Q_L + Q_H$; $\Delta_H(z^*)$ and $\Delta_L(z^*)$ represent the average quality of sellers above and below the threshold, respectively; and $m_H(z^*)$ and $m_L(z^*)$ represent of share of the entrant cohort above and below the threshold, respectively.

We are interested in the comparative statics of making the badge more restrictive by increasing

z^* from z_2 to z_3 and only awarding it to the highest-quality sellers. For ease of notation, let p_H^1, p_L^1 be the prices under $z^* = z_2$, and p_H^2, p_L^2 be the prices under $z^* = z_3$.

Lemma 1 . $p_L^2 < p_H^1$.

Proof. Suppose instead that $p_L^2 \geq p_H^1$. Since $p_H^2 > p_L^2$, it follows that both prices have increased. Hence the total output must increase too (i.e., $Q_2 > Q_1$). Then $p_L^2 = P(Q_2) + \Delta_L(z_3) < P(Q_1) + \Delta_H(z_2) = p_H^1$, which is a contradiction. ■

The above Lemma shows that the transition will hurt the middle-quality sellers who lose the badge.[3] This results in a lower fixed-cost entry threshold for these sellers and fewer will enter. The effect on the other two groups depends on the parameters of the model, such as marginal cost, entry cost, and quality levels. However, we can show that at least one of the two prices should go up.

Proposition 1 *At least one price will increase under $z^* = z_3$.*

Proof. Assume instead that both $p_L^2 < p_L^1$ and $p_H^2 < p_H^1$. Because both $q(p)$ and $n(p)$ are increasing in p, it follows that

$$\begin{aligned}
Q_2 &= q\left(p_H^2\right) n\left(p_H^2\right) m_H(z_3) + q\left(p_L^2\right) n\left(p_L^2\right) m_L(z_3) \\
&= q\left(p_H^2\right) n\left(p_H^2\right) m_3 + q\left(p_L^2\right) n\left(p_L^2\right)(m_1 + m_2) \\
&< q\left(p_H^1\right) n\left(p_H^1\right) m_3 + q\left(p_L^1\right) n\left(p_L^1\right)(m_1 + m_2) \\
&= q\left(p_H^1\right) n\left(p_H^1\right)(m_3 + m_2) + q\left(p_L^1\right) n\left(p_L^1\right)(m_1) \\
&\quad - m_2(q\left(p_H^1\right) n\left(p_H^1\right) - q\left(p_L^1\right) n\left(p_L^1\right)) \\
&< Q_1.
\end{aligned}$$

But if total output decreases, and both quality premiums increase, then both prices must increase, yielding a contradiction. ■

The increase in the price and distribution of the fixed costs determine the size and quality of sellers in the market. Since at least one price must increase, causing variable profits to increase and offset higher fixed cost of entry, then if $p_H^2 > p_H^1$, more sellers of the highest quality z_3 will enter the market. Similarly, if $p_L^2 > p_L^1$ then sellers of the lowest quality z_1 will have higher incentives

[3]The proof requires a bit more than the trivial convex combination of quality levels because changes in price affect quantity, and that feedbacks into both prices.

to enter the market as well. We use this comparative statistic to guide our empirical analysis. A straightforward extension in which one group of sellers has uniformly higher entry cost, the sellers who enter the market after an increase in the price will have a higher quality level on average. We show evidence of this in Section 7.2.

Two final notes are warranted. First, our simple framework assumes that quality is fixed as in a standard adverse selection model. Sellers cannot exert effort to improve their quality as in a typical moral hazard model. We cannot of course measure effort or lack thereof directly. However, in Section 6.4 we show evidence consistent with the simple framework we use.

Second, unlike our comparative static analysis, actual markets will adjust over time so that entrants will change the composition of badged and non-badged sellers, leading prices and entry rates to adjust to a new equilibrium. We explore this by considering different time frames in our empirical analyses, as well as considering the incentives to enter. Namely, if badges become a lot more scarce, then badges will command a premium, and those sellers who believe they have a shot as getting a badge will be more eager to enter. We discuss this in more detail in Section 5.

4 Data

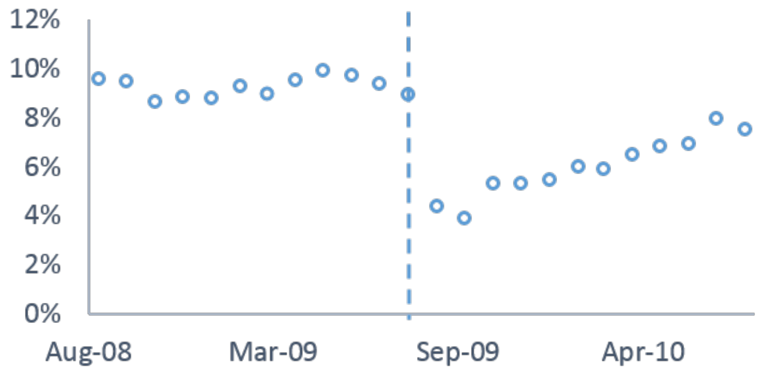
We use proprietary data from eBay that include detailed characteristics on product attributes, listing features, buyer history, and seller feedback and reputation. We begin with data from October 2008 to September 2010, which include all listing and transaction data in the year before and the year after the policy change.

One important feature of our data is information on product subcategories cataloged by eBay. There are about 400 subcategories, such as Office, Lamps and Lighting, Beads and Jewelry Making, Video Game Memorabilia, Digital Cameras, Makeup, and many others. A subcategory is the finest level of eBay's catalog that includes all listings on the site.[4]

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

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

Figure 1: Share of Badged Sellers



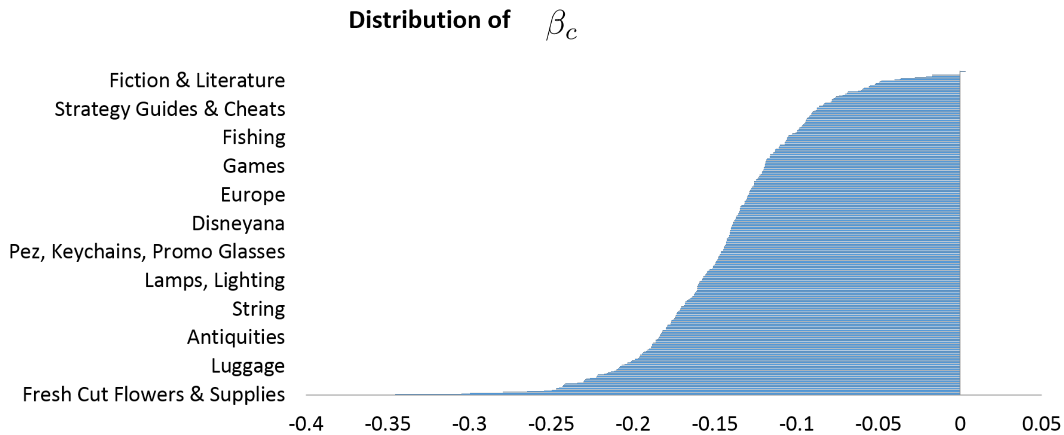
Finally, the use of internal data allows us to construct a measure of quality that is not observed with information that appears on the page. Every seller has a reputation score and percent-positive (PP) on ebay, the latter being the number of positive ratings divided by the total number of ratings. Nosko and Tadelis [2015] demonstrate the extreme skewness of PP, where the mean is 99.3% ad the median is 100%. They conjecture that silence is itself a sign of some negativity, and they construct a new measure that they call "Effective Percentage Positive (EPP), which is the number of positive feedback transactions divided by the number of total transactions. Nosko and Tadelis [2015] have shown that EPP contains much more information on transaction quality than conventional feedback and reputation scores. We follow their approach and for each seller we compute its EPP and use it as a measure of quality.

5 Empirical Strategy

We use the policy change described in section 2 as a quasi-experiment. Figure 1 clearly shows the policy change caused a significant decrease in the share and number of badged sellers. The average share of badged sellers was around 10% throughout the year before the policy change, and dropped to 4% right after the policy change, with some adjustment taking place in the following year.

Our goal is to find what can be considered exogenous variation in the share of badged sellers across different subcategories because of the differential ways in which the policy affects different subcategories. To compute the change in the share of badged sellers across different subcategories, we could use the immediate change in this measure in the week before and the week after the policy change. However, this approach would neglect any time trends. To account for the trend, we estimate the change in the share of badged sellers in different subcategories using the following

Figure 2: Heterogeneous Impact of Policy Change on Different Subcategories



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.

event study approach:

$$Share\_Badged_{ct} = \beta_c Policy + \eta_c + \alpha_c t + \epsilon_{ct}, \tag{1}$$

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; η_c are subcategory fixed effects; α_c is a subcategory-specific linear time trend; and ϵ_{ct} are error terms.

The horizontal bars in Figure 2 are the estimated changes in the share of badged sellers caused by the policy change using a period of six months before and six months after the policy change. The figure shows that the decrease in the share of badged sellers after the policy change is very different across different markets and for some subcategories is as large as 35%. This large variation in the share of badged sellers after the policy across subcategories is robust to different specifications.[5]

Our identification strategy exploits the variability in estimated changes in the share of badged sellers ($\widehat{\beta_c}$) in different subcategories induced by the policy change to identify the impact on the number and quality of entrants using a continuous difference-in-difference (DiD) approach. In particular, we estimate the policy impact by comparing the intertemporal changes in the number and quality of entrants in the subcategories that are more affected by the policy change against

[5] As robustness tests, we also tried estimating the same equation considering a period of one, three, and twelve months before and after the policy change, and the qualitative results are similar. We have also directly compute the drop in share of badged sellers using the average share of badged sellers in the week before and the week after the policy change in different markets, and the results are similar. In addition to using the absolute value of estimated changes in the share of badged sellers across subcategories, we also use the percentiles of these estimates across subcategories, namely $\widehat{\beta_c^{pct}}$ in place of $\widehat{\beta_c}$, as robustness checks for our second-stage regression.

intertemporal changes of these two measures in the subcategories that are less affected over the same time periods. This DiD approach is continuous in the sense that the "treatments" (i.e., policy impacts on the share of badged sellers across subcategories) take continuous values between 0 and 1. Specifically, the DiD specification is given as

$$Y_{ct} = \gamma \widehat{\beta}_c Policy + \mu_c + \xi_t + \epsilon_{ct}, \tag{2}$$

where Y_{ct} are the outcome variables of interest in subcategory c in month t (e.g., quality, or entry); $\widehat{\beta}_c$ is the estimated policy impact on the share of badged sellers from our first stage (1) shown in Figure 2; μ_c are subcategory fixed effects; ξ_t are month fixed effects; and ϵ_{ct} are error terms.

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

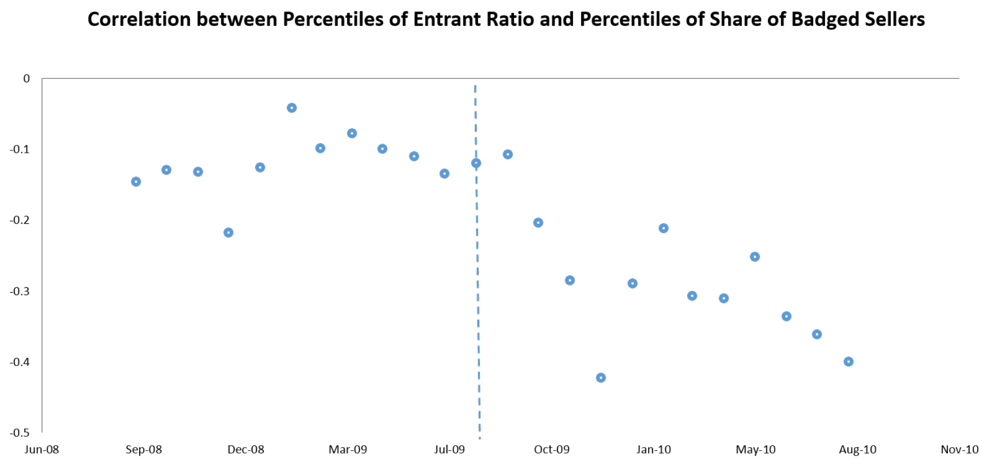
Note that we have two types of entrants: new sellers on eBay (15%) and existing sellers entering new subcategories (85%). An implication of our theoretical framework is that these two types of entry may behave differently if they differ in their entry costs, which is a reasonable assumption. In our main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

The DiD approach controls for time-invariant differences in the variables of interest across subcategories; for example, the entrant ratio in the Clothing market is higher than that in the Antiques market. The approach also controls for differences in the entrant ratio over time, for example, changes in the overall popularity of selling on eBay over time. As in most DiD approaches, our key identification assumption for a causal interpretation of $\widehat{\gamma}$ is that serially correlated unobserved errors do not systematically correlate with $\widehat{\beta}_c$ and Y_{ct} simultaneously. We provide a robustness test of this identification assumption in Table 4 in Section 7.

6 Results

We first estimate the effects of the policy change on the average rate of entry and quality provided by the entrants. We then describe how the quality distribution of the entrants changes. Subsequently, we investigate whether the changes in quality are more likely to be due to changes in sellers selection

Figure 3: Market Structure and Entry



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

or behavior. Finally, we study how prices, sales probability, and market share for different groups of sellers change.

6.1 Effect on Number of Entrants

Figure 2 shows that the policy change resulted in a great deal of heterogeneity in changes in the share of badged sellers across different subcategories. Our theoretical benchmark suggests that variation in how hard it is to obtain a badge will impact both the number and quality of entrants. Before conducting the two-stage regression analysis described above, we first provide descriptive evidence that the heterogeneity in policy impact has meaningful implications for the rate of entry. To normalize across subcategories, we define the entrant ratio to be the number of entrants in month t divided by the number of sellers in month $t-1$ in a particular category. Figure 3 plots the correlation between the entrant ratio and the share of badged sellers in each subcategory.

Given that both the entrant ratio and the share of badged sellers responded to the policy change, we proceed with some simple descriptive facts by normalizing these two measures for a meaningful comparison. In particular, we compute the percentiles of these two measures across subcategories and plot their correlation. Figure 3 shows that there is a negative correlation between the entrant ratio and the share of badged sellers across subcategories, i.e., subcategories with a larger share of badged sellers are associated with smaller entrant ratios. This correlation becomes more negative after the policy change, marked by the dashed vertical line, suggesting that the policy change affected the entry pattern in different subcategories, and the magnitude is correlated with changes in the share of badged sellers.

Table 1 reports $\widehat{\gamma}$ from regression (2) for six variables measuring entry, each in a separate panel. Recall that since β_c estimates in Figure 2 are negative, a negative γ implies an increase in the outcome variable in subcategories that are more affected by the policy change. Panel A column 1 shows that the entrant ratio is higher in subcategories that are more affected, using data from three months before and after the policy change. In particular, a 10% larger decrease in the share of badged sellers leads to 3% 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, [6] where the estimate is smaller and is not statistically significant. This suggests that the market stabilizes on a new equilibrium after the first six months.

To understand the distributional impact of the policy change on the number of entrants, in Figure 4a we plot two time series, monthly average (normalized) number of entrants and monthly average share of badged sellers, in the subcategories that are most affected (top quintile of β_c) and least affected (bottom 20 quintile of β_c), respectively. Figure 4a shows that in the top 20 percentile of the subcategories, the share of badged sellers decreases from about 35% to less than 15% right after the policy change, whereas in the bottom 20 percentile, the share of badged sellers decreases from about 18% to 10%. On the other hand, the average monthly number of entrants in the top 20 percentile increases by 25%, whereas there is no obvious change in the number of entrants for the bottom 20 percentile of subcategories. This suggests that the policy effect on entry comes mainly from subcategories that were heavily affected. Additionally, the entry rates seem to stabilize after three months. We show the robustness of these results by looking at top and bottom deciles of β_c in Figure 9.

6.2 Performance of the Entrant Cohort

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

We contruct a seller's EPP using the number of transactions and the number of positive feedback

[6] We do not include longer time periods because eBay has made other significant changes to its trust systems.

Table 1: Policy Impact on Rate and Quality of Entrants

Panel A. Entrant Ratio			
	(1)	(2)	(3)
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-0.299***	-0.204***	-0.047
	(0.041)	(0.027)	(0.051)
R^2	0.913	0.889	0.691
Panel B. EPP Conditional on Survival in the Second Year			
	+/- 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
Panel C. Sales Quantity Conditional on Survival in the Second Year			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	15.082***	2.867	2.560
	(4.455)	(2.877)	(3.533)
R^2	0.605	0.549	0.505
Panel D. Total Sales			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	-6883	-9895***	-4737
	(6611)	(4025)	(3678)
R^2	0.930	0.930	0.942
Panel E. 2nd-yr Sales Quantity/ # Entrants			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	5.573**	2.002	3.046
	(0.039)	(2.042)	(2.121)
R^2	0.496	0.404	0.381
Panel F. 2nd-yr Sales Quantity			
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	1098	-3015	11644
	(10375)	(6801)	(7477)
R^2	0.745	0.736	0.723

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

*** indicates significance at p = 0.01; ** indicates p = 0.05.

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

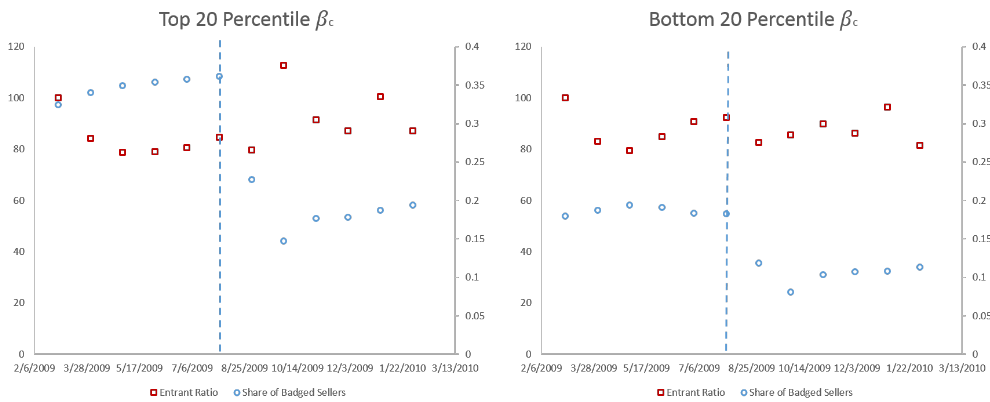
Negative coefficients in Panel B in Table 1 show that there is an increase in the average average quality of entrants in the more affected subcategories after the policy change. This effect stabilizes from -10% to -6.6% as we expand the window length from six to twelve months. Column 3 shows that the 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 4b we show the average EPP for entrants in the top and bottom quintiles of the affected subcategories. Note that EPP is decreasing on eBay over time because buyers are less likely to leave feedback in general, but the average EPP is higher for the top quintile of the affected subcategories compared to the bottom quintile.

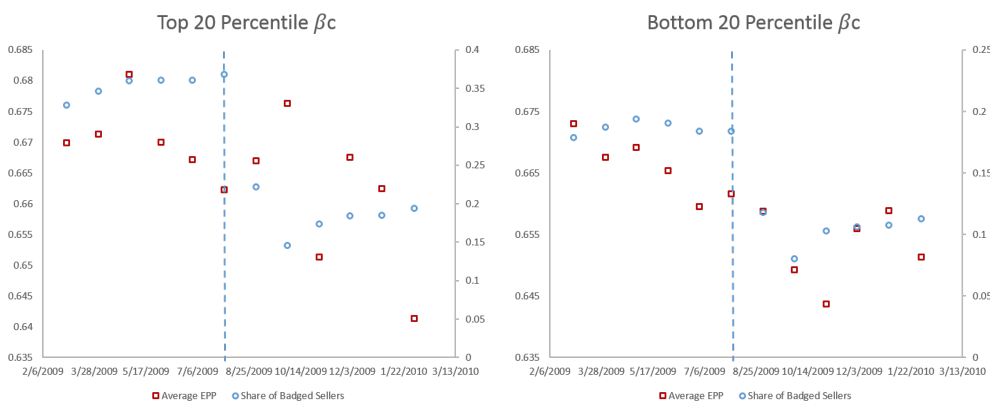
Next, we look at the average and total sales quantity during a seller's first year of entry, conditional on it surviving in the second year. A positive and significant coefficient in column 1 of Panel C shows that over the short term, the sales quantity from each entrants is smaller in subcategories affected more by the policy change; however, this drop becomes insignificant when considering a longer time period. Additionally, from the seventh to the twelfth month after the policy change, the change in entrants' sales quantity remains statistically insignificant. This result indicates that the average entrant is smaller in the subcategories most affected by the policy change. Recall that these subcategories have more entrants on average as well. As a result, this regression does not necessarily imply a decrease in the total number of sales by entrants. In fact, when we run a regression of total sales by entrants in Panel D, we observe that the subcategories more affected by the policy change have a higher total number of sales by entrants. Additionally, we plot analogous graphs to analyze the distributional policy impact on entrants' sales by the top and bottom quintiles of the affected subcategories in Figure 4c, which shows a short-run surge in the number of total sales in the top quintile of the affected subcategories with very little impact on the bottom quintile of the affected subcategories.

Finally, we study entrants' survival by looking at the average size of entrants in the year after entry, assigning 0 to sellers who do not sell any items in their second year. The advantage of this measure over a simple survival dummy is that it is able to capture a seller's change in size as well

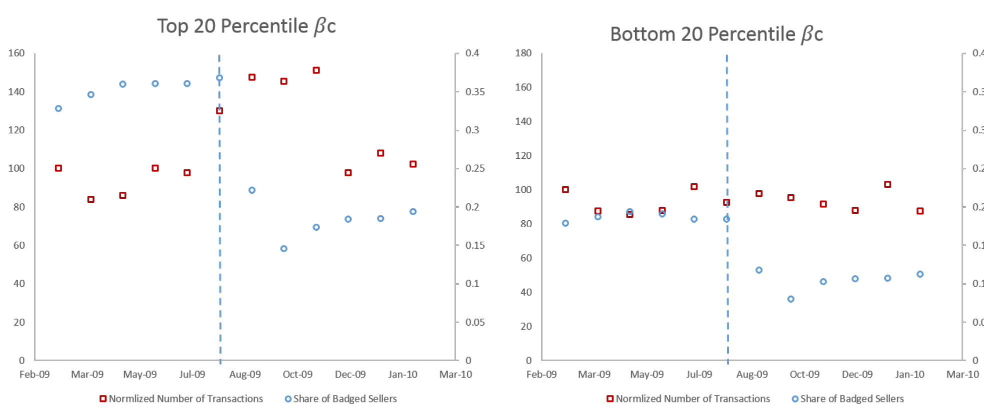
Figure 4: Distributional Policy Impact on Entrants



(a) Distributional Policy Impact on Number of Entrants



(b) Distributional Policy Impact on EPP



(c) Distributional Policy Impact on Sales

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

as exit.[7] Panel E shows that the average sales quantity in the second year per entrant decreases more for entrants in subcategories that are more affected by the policy change. This observation is consistent with entrants being smaller in the affected area, as shown in Panel C. However, this effect becomes insignificant when we consider a period of seven to twelve months after the policy change. Additionally, the total number of items sold by entrants in the second year does not change significantly, as shown in Panel F.

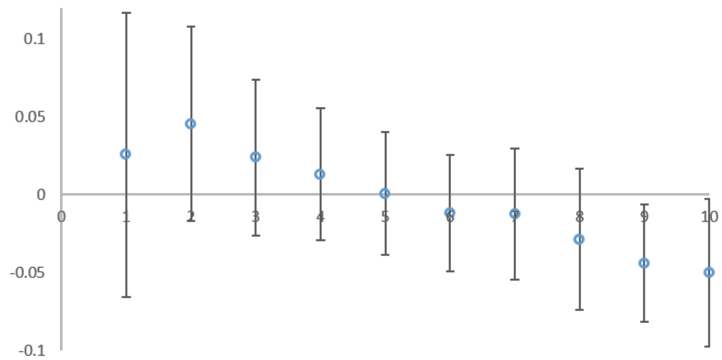
6.3 Quality Distribution of the Entrant Cohort

An important implication of our simple theoretical framework is how the distribution of entrants' quality varies after the policy change. Our theory predicts that under mild assumptions, there should be more entrants of high quality, as the benefit of getting a more selective and informative badge is higher. Additionally, the theory predicts that low-quality sellers may enter more often because they are pooling with a better set of sellers who lost their badge, implying higher average prices and/or sales for unbadged sellers in equilibrium. To test this prediction, within each subcategory, we partition entrants into deciles based on their EPP score in the first year after their entry. For example, we look at entrants within the top decile as determined by their EPP score based on their transactions in the first year after their entry. Then we perform the DiD specification for this decile and check if these EPPs have increased more for the subcategories more affected by the policy change. A negative number will indicate a fatter tail of the high-end of the distribution. Respectively, if we look at the bottom decile of entrants in terms of their EPP and compare across subcategories, a positive estimate means a fatter tail on the low end. Another prediction was that the sellers who had a chance of becoming badged before and no longer have this opportunity after the policy change will enter less often. A distribution of entrants' quality with a fatter tail from both left and right will indicate a smaller share of average-quality entrants.

We plot the change in first-year EPP for entrants of different quality deciles in Figure 5. For consistency, we condition the EPP calculation on an entrant's survival in the second year. Entrants are counted every two months. To be able to take the average of cohorts, we restrict our attention to subcategories with at least 100 entrants. As a result, for each subcategory, we have three observations (six-month equivalent) before the policy change and three observations after it. Additionally, we only consider subcategories that have entry in all of the six two-month periods

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

Figure 5: Change in EPP for Entrants in Different Quality Deciles



Notes: Bars indicate 95% confidence intervals.

and remove subcategories with a small number of entrants. This leaves us with 228 out of the 400 subcategories.

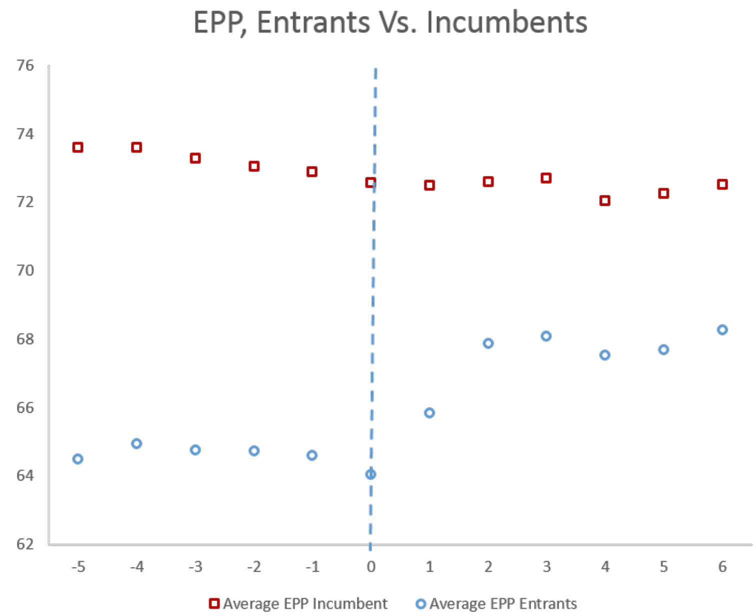
The x-axis in Figure 5 indicates different quality deciles, with "10" being the highest decile of EPP and "1" being the lowest decile of EPP. The figure plots point estimates of the changes in EPP for the entrant cohorts with 95% confidence intervals. The top-two decile point estimates are statistically negative, as predicted. Though the other estimates are not significant from zero, we do observe a monotonically decreasing relationship that is consistent with our prediction that the quality distribution of entrants after the policy change varies and has fatter tails because sellers from the extremes of the quality distribution now have stronger incentives to enter. This in turn implies that sellers in the middle of the quality distribution enter less frequently.

6.4 Impact on Incumbents

In this section, we study how the EPP of incumbents has changed. This exercise is interesting to understand how incumbents respond to an increase in the certification threshold. In particular, the results in the previous subsections show that the policy change had an impact on the entry decision of sellers into different subcategories, and that this impact differs among entrants of different quality levels, suggesting a selection-of-entrants interpretation. However, the impact on the quality provided by entrants could in principle be solely driven by a moral-hazard story, where similar entrants changed their behavior after the policy change depending on the subcategory they entered.

Figure 6 plots the average monthly EPP of all entrants and incumbents within six months of the policy change. Incumbents are defined as sellers who listed at least one item before and one

Figure 6: Change in EPP of Incumbents and Entrants



item after the change. As indicated in the figure, the average EPP of entrants has increased, while no apparent change in the average EPP of incumbents is observed.

To add more rigor, we perform our DiD analyses for incumbents only. We define incumbents to be sellers who listed at least one item before and one item after the policy change in the specified time windows. In Panel A of Table 2, we see that estimates are either borderline significant or non-significant. To test the robustness of these estimates, we use percentiles of β_c across subcategories, rather than using the absolute values, for the DiD analyses. Under the alternative specification, none of the estimates for the three time windows is statistically significant at the 10% level. In addition, we note that the point estimate in the alternative specification yields the opposite interpretation as the estimate (0.044) in column 1, suggesting that even this borderline significant result is not robust. We also study the distributional impact of the policy on EPP and find no difference in its impact for the top and bottom quintiles of subcategories. This evidence shows that the EPP of incumbents did not increase after the policy change.

Having established that there is little change in average EPP among all incumbents, we next study whether there is any change in some incumbents' quality. We first repeat the DiD analyses for sellers who entered not too early before the policy change. The idea is that these sellers could be similar to those that entered right after the policy change because of their proximity in entry date. In Panel B of Table 2, we study how EPP changes for sellers that entered either three months

Table 2: Policy Impact on Quality of Incumbents

Panel A. EPP from Incumbents			
	(1)	(2)	(3)
	+/- 3 Months	+/- 6 Months	Month 7 to 12
Estimate	0.044*	-0.020	0.021
	(0.025)	(0.018)	(0.021)
R^2	0.887	0.853	0.823
Panel B. Sellers who Entered n Months before the Policy			
	n=3	n=6	
Estimate	0.068	-0.027	
	(0.054)	(0.053)	
R^2	0.459	0.415	

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 = 0.01; ** indicates p = 0.05.

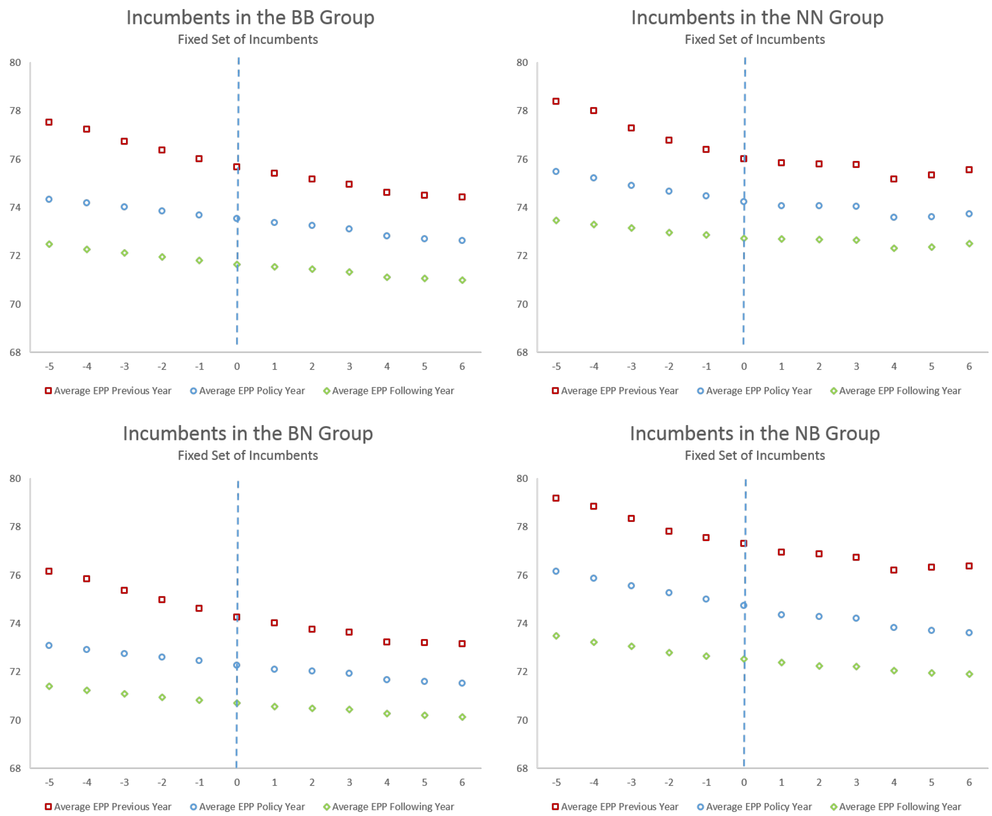
or six months before the policy change. The insignificant estimates show that there are no changes in behavior for these two groups of sellers, suggesting that a significant share of the changes in EPP from entrants is likely to come from improved selection.

Any incumbent can be badged or not badged before and after the policy change. Hence, we can divide incumbents into four collectively exhaustive partitions 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. Because we consider data from up to a year before and after the change, we define a seller as belonging to the BB group if it is a Powerseller for at least eleven out of twelve of its active months on eBay before the policy change (91.6% of the time) and is eTRS for at least 91.6% of its active months after the policy change. If this requirement is not met the seller is categorized as N in the relevant period. Note that sellers in all these four groups must be active (listed at least one item) both before and after the policy change.[8] The largest group is the NN group with over 50% of sellers, while the NB group is the smallest at 4%.

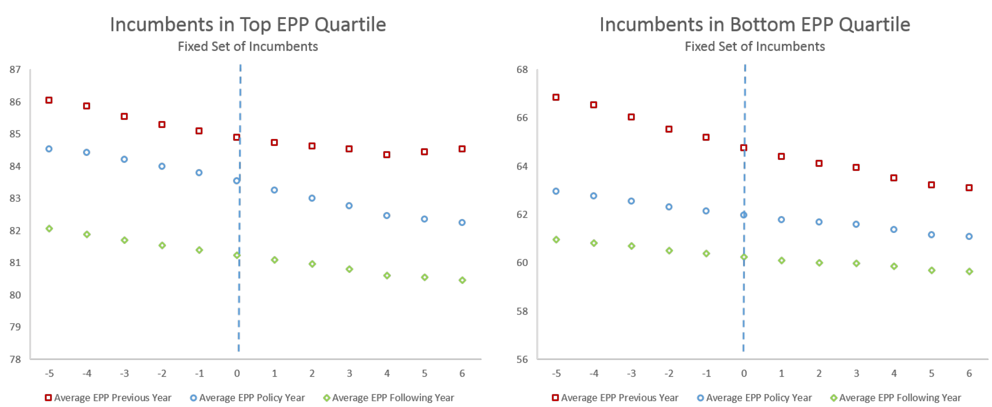
In Figure 7a, we plot the average monthly EPP for incumbents in the four groups. The solid line is the average monthly EPP provided by incumbents from a particular group in the six months before and after the policy implementation date. The dotted line and dashed-dotted line are the

[8] For robustness, we also change the sample period from twelve months to six and three months, and change the threshold for each group to 75% and 50%, respectively. The results are qualitatively similar.

Figure 7: Change in EPP of Incumbents



(a) Four Groups of Incumbents



(b) Top Vs. Bottom 20 Percentile

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.

average EPP provided by the same set of incumbents in the same months in the previous year and the following year, respectively. We see that there is no obvious difference between the EPP provided by incumbents in the year of the policy change and the EPP in adjacent years, except that EPPs are getting lower over time.[9] This implies that the change in the average monthly EPP observed in these two figures is due to seasonality.

We created a similar plot for sellers of different quality quartiles measured by EPP. 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.

Disentangling improved selection from better behavior is tricky for entrants. The reason is that, unlike in the case of incumbents, we cannot fix a set of entrants and track their behavior before they enter the market. However, we believe that the lack of change in incumbents' behavior overall and in different partitions strongly suggests that a significant fraction of the increase in quality provided by entrants at the tails of the quality distribution is likely due to selection rather than to behavioral changes.

6.5 Impact on Badge Premium

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

Following the literature that studies price changes on eBay (e.g., Einav et al. [2011] and Hui et al. [2016]), we take advantage of product ID's in our data to construct an average price for each product that was listed as fixed-price and sold. For each individual item sold we define its "relative price" as the item's price divided by the average price of the product. In column 1 of table 3, we study the change in the price premium, which is the relative prices, for different groups of sellers using transactions from one month before and one month after the policy change, where NN is the excluded group. We find that the sellers in the BN group experience a statistically significant decrease in relative price of 0.7% (relative to the insignificant 0.3% decrease in the NN group).

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

Table 3: Change in Badge Premium

	(1) Relative Price	(2) Sales Probability	(3) Sales Quantity	(4) Market Share
Policy	-0.003	0.015***	0.009	-1.5E-07(-2%)
	(0.003)	(0.001)	(0.006)	(1.4E-06)
BB*Policy	-0.003	0.024***	0.032***	6.2E-06***(15%)
	(0.003)	(0.001)	(0.005)	(2.2E-06)
BN*Policy	-0.007***	-0.001***	-0.010***	-3.3E-06*(-6%)
	(0.002)	(4.E-04)	(0.004)	(1.8E-06)
NB*Policy	0.001	0.097***	0.221***	1.8E-06(13%)
	(0.012)	(0.003)	(0.026)	(4.1E-06)
Seller FE	✓	✓	✓	✓
Week FE	✓	✓	✓	✓
R^2	0.288	0.808	0.862	0.813

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 = 0.01; ** indicates p = 0.05; * indicates p = 0.1.

The changes in relative price for the other groups are not statistically significant. 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.

In columns 2 and 3, we show the changes in badge premium in terms of sales probability and sales quantity using transactions from one month before and one month after the policy change. We see that all groups of sellers except for group BN experience an increase in both measures. The magnitudes for both measures in descending order are NB, BB, NN, and BN. Our interpretation is that the sellers in the NB group experience an increase in sales because they gain the reputation badge.[10] The sellers in the BB group experience an increase in sales because the new badge conveys more information and therefore is more valuable than the old one. The sellers in the NN group are better off because they are being pooled with higher-quality sellers than before. Finally, the sellers in the BN group are worse off because they lose their badge. Combining the estimates for the BN group in columns 1 and 3, 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

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

theory that middle-quality sellers are hurt by the policy change.

Finally, we analyze the policy impact on market share for different groups of sellers using their transactions from one month before and one month after the policy change. This regression is at the seller–week level so that the market share of a seller in a given week equals the number of transactions of that seller divided by the total number of transactions in that week. If a seller does not have any sale in a particular week, we fill in zero as that seller's market share for that week. We report the estimates in column 4 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 15% as well because the change in market share for the NN group is small. On the other hand, the BN group had a 6% smaller (relative and net) market share after the policy change, although the result is not as significant. We also performed the same set of regressions using transactions in the three months before and three months after the policy change. The results are similar and reported in table 7.

Consider all the estimates in Table 3 together, we see that after the policy change the BN group is worse off and the other three groups are better off mostly through increased sales.

7 Robustness Checks

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

7.1 Placebo Test on the Exclusion Restriction

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

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

We test this using a placebo test: we use the $\widehat{\beta_c}$ estimated from the year of the policy change, and re-perform the second-stage regression using data around September in the previous year. In Table 4, we report the estimated γ for entrant ratio, EPP, and total sales for entrants in the previous year. Neither of the estimates is statistically significant in this table, implying that the impact of the policy change on the share of badged sellers in different subcategories is as good as random with respect to different entry variables across subcategories in the previous year. This suggests that the policy change generates some exogenous variations in share of badged sellers across subcategories that are not mere artifacts of heterogeneities across subcategories.

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

7.2 Two Types of Entry

We distinguish between two types of entrants into a subcategory: new sellers on eBay and existing sellers entering a new subcategory. From the lens of our theoretical model, these two types of entrants differ in their entry cost: the entry cost of starting to sell on eBay must be higher than the entry cost of selling in new subcategories for existing eBay sellers who are already familiar with the platform. In practice, it is possible that sellers do not make entry decisions into different subcategories based on changes in price in that category alone. Rather, they may have reasons to sell in more than one category, such as meeting both the value and quantity requirement for the eTRS badge that is not category-specific. For example, a laptop seller might find it harder to meet the minimum quantity requirement for the new badge, and therefore enters a cheap category like cables merely to meet this requirement.

We find that among entrants into new markets, about 15% of them are new sellers on the platform and 85% are existing sellers entering new subcategories. Next, we perform our previous DiD analyses for the two types separately (see Table 5). Figure 10 shows the change in EPP deciles similar to Figure 5. Results in both exercises are very similar across the two types, and the relative magnitudes of these estimates are consistent with our theory. Assuming that entry costs of starting to sell on eBay are higher than those of entering a new category for an existing seller on eBay, new sellers need to have higher quality to compensate for the entry cost relative to the increase in quality among existing sellers. By the same logic, there should be more entry of the existing sellers relative to the increase in entry of new sellers.

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

7.3 Rate and Quality of Entry

In this section, we test the robustness of our results in Table 1. We replicate the DiD estimation using percentiles of $\widehat{\beta_c}$, rather than their actual values. This normalization enables scale-free comparisons across subcategories. Note that the signs of the estimates should flip because the actual change in share of badged sellers was negative, and a high percentile of $\widehat{\beta_c}$ corresponds to a more negative $\widehat{\beta_c}$.

In Panel A of Table 6, we see that the signs are indeed flipped compared to Panel A in Table 1. We also observe the same pattern, that the impact of the policy change on the entrant ratio gets smaller over time, and becomes insignificant for months seven to twelve after the policy change. In Panel B, we see that, consistent with our previous finding, the average EPP is higher in the cateogeries that are more affected. In fact, all estimates in Panel C through Panel F have the opposite signs as their counterparts in Table 1. The consistency in qualitative results across the two measures of β_c is reassuring.

Subsequently, we test whether our finding on the distributional policy impact holds if we look at the top and bottom 10 percentiles of affected subcategories. Figure 9 plots a parallel to Figure 4 with the top and bottom 10 percentiles. The qualitative results still hold, that is, most responses in entrant ratios and EPP come from subcategories that are most affected. We have also tried removing outliers and the result remains the same.

Finally, we consider different window lengths for defining EPP for incumbents. In Figure 10, EPP is defined over transactions of a seller within the past six months of any given month. As shown, there do not seem to be changes in the average quality of incumbents, but there is an increase in the average quality of entrants immediately after the policy change. We also tried using the past three months and one month in defining incumbents' EPP, as well as using the future six months, one year, and three years, and the results are similar.

7.4 Badge Premium

In Table 7, we perform the same set of regressions as in Table 3 except that we now use data from three months before and three months after the policy change for estimation. We see that the relative changes across different groups carry through in this analysis.

8 Conclusion

Following a policy change on eBay, more demanding criteria and more precise information are used to award a quality-signalling badge to sellers. We use this change to gauge insight into the effects of more stringent certification and reputation measures on entry, which is a hard-to-study relationship. We exploit the differential impact of the policy change on different subcategories of sellers for identification, and document a negative correlation between the share of badged sellers and the rate of entry across subcategories affected by the policy change. The subcategories that experience a higher reduction in the share of badged sellers because of the policy change have larger entry rates after the policy change. However, this effect is temporary, and tends to disappear once the market adjusts to the new equilibrium, after about six months.

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

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

References

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

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

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

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

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

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

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

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

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

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

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

Hugo Hopenhayn and Maryam Saeedi. Reputation signals and market outcomes. Technical report, National Bureau of Economic Research, 2016.

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

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

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

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

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

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

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