

Market Structure, Reputation, and the Value of Quality Certification[†]

By DANIEL W. ELFENBEIN, RAYMOND FISMAN, AND BRIAN MCMANUS*

Quality certification programs help consumers identify high-quality products or sellers in markets with information asymmetries. Using data from eBay UK's online marketplace, we study how certification's impact on demand varies with market- and seller-level attributes, exploiting variation in sellers' certification status within groups of near-identical listings. The positive effects of eBay's "top rated seller" certification are stronger for categories with few other certified sellers, in more competitive markets, and for sellers with shorter records of past performance. These findings indicate certification provides more value when certification is rare, the product space is crowded, and for sellers lacking established reputations. (JEL D12, D82, L15, L86)

Quality certification as a means of reducing “lemons” problems is common across a range of markets, including consumer retail, firm-to-firm trade, labor markets, and specialty services like medical treatment and auto repair. The ubiquity of “expert” guidance from certifiers like Consumer Reports, city hygiene departments, and hospital credentialing organizations, to name a few, has spawned a large empirical literature (surveyed by Dranove and Jin 2010) that has provided many useful insights on the efficacy and attributes of certification mechanisms.

Empirical studies in this area generally feature a close examination of how quality certification affects trade in a single market, where the competitive environment and information context is held fixed. This makes it difficult to answer a host of fundamental questions, such as: How does the number of competing products on nearby store shelves or web pages affect consumer responses to certification? How might certification's value differ between industries with a high scope for seller opportunism—as in automotive services—versus those with little quality uncertainty—like the market for toasters? And how do alternative reputational signals like crowd-sourced feedback affect the value of certification? Developing a better understanding of the

*Elfenbein: Olin Business School, Washington University in St. Louis, Campus Box 1156, One Brookings Drive, St. Louis, MO 63130 (e-mail: elfenbein@wustl.edu); Fisman: Boston University Department of Economics, 270 Bay State Road, Boston, MA 02215 and National Bureau of Economic Research (NBER) (e-mail: rfisman@bu.edu); McManus: University of North Carolina, Chapel Hill, Department of Economics CB 3305, Chapel Hill, NC 27599 (e-mail: mcmanusb@unc.edu). We thank eBay, and in particular Brad Matthews, Jana Moles, Jana Honnerova, Irene Garcia Sacedon, and Vasanth Yenegalla. Additionally, we thank seminar and conference participants at Arizona State University, the Federal Trade Commission, Georgia Tech, Northwestern University's Searle Center, Oxford University's Centre for Corporate Reputation, UCLA, the University of North Carolina–Charlotte, the University of Michigan, and the University of Minnesota for their helpful comments.

[†]Go to <http://dx.doi.org/10.1257/mic.20130182> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

differential impact of certification as a function of seller or industry attributes can help to assess how the introduction of certification mechanisms is likely to affect industry evolution, including firm entry and market power, and to anticipate what types of firms will gain and lose as a result. Evaluating the heterogeneous effects of certification may also aid in the design of certification programs themselves, providing guidance, for example, on the choice of quality threshold requirements and how the program might be influenced by industry attributes like the scope for opportunism.

To begin addressing these questions, we examine eBay's Top-Rated Seller (eTRS) certification using a dataset comprised of 1.6 million groups of near-identical matched listings posted by nearly 23,000 sellers in more than 8,000 distinct product categories in the United Kingdom. Groups of matched listings often feature variation in sellers' certification status for otherwise identical products. The same criteria are used to evaluate all sellers for certification, independent of the category where they place their listings. Heterogeneity across product categories and sellers allows us to examine how consumer responses to certification vary across markets that differ in concentration and in underlying quality uncertainty, and how certification's impact varies with the sellers' scale and reputation.

Institutional features of the eTRS program—especially the fact that sellers gain and lose eTRS certification abruptly—allow us to cleanly identify the impact of quality certification on demand by isolating cases where sellers list identical items at the same price with and without certification. This enables us to capture the impact of certification on demand, holding price (and other attributes) constant. Additionally, our data include information about the number of times a listing is served to consumers as a search result and also how frequently a listing's detailed web page is displayed so we may further separate the “advertising” impact of certification from its informational content regarding seller quality. (In most contexts, it is difficult to make this distinction; for example, Consumer Reports puts top-ranked items first, and may not even list lesser quality products.)

A priori, it is unclear how differences in market attributes may affect the impact of certification on consumer demand. If transaction risk, conditional on observables, is identical across sellers then certification's premium could be independent of the proportion of certified sellers in the marketplace. Alternatively, if a scarcity of certified sellers is interpreted by consumers as a sign of market-specific quality concerns, certification's value may be decreasing in the proportion of certified sellers. There is similar ambiguity in the case of market concentration. Certification may be an increasingly useful means of “vertical” differentiation as horizontal product market competition intensifies; alternatively consumers may view the additional benefits to a particular product's certification as having a fixed value, unrelated to the number of other products in a category. Our findings offer some guidance in assessing which of these effects dominate empirically.

To provide a conceptual underpinning for the empirical regularities we observe, we present a simple theoretical model based on unobserved seller quality, imperfect customer feedback, and the assumption that eBay has an information advantage relative to buyers. The model highlights an intuitive mechanism relating eTRS prevalence to the certification premium. In product categories with lower average seller reliability, fewer firms obtain certification and the value of certification is

thus greater since it carries more informational content. Modeling competition as horizontal product differentiation, we also show that eTRS status will have a greater impact on sales probability in more competitive categories (holding price constant), since vertical differentiation is more valuable in the face of intense product market competition. The model also generates the straightforward prediction that the value of outside certification becomes less important (for high-quality sellers) as the seller accumulates publicly visible customer feedback.

We begin by documenting that gaining eTRS certification—displayed to consumers via a small eTRS “badge” on all of a seller’s listings—is associated with a significant increase in demand even when controlling for the impact of certification on the listing’s prominence in search results. Certification increases by 7 percent the probability that a given seller’s fixed-price listing ends successfully with a sale. Combined with price sensitivity estimates, our findings imply that consumers will pay about 7 percent more for certified items.

We then proceed to analyze how market and seller attributes affect the value of certification. First, we show that in categories with few other badged listings—where we argue that concerns of unobserved seller quality are most prevalent—the impact of certification is much greater. For listings in categories at the twenty-fifth percentile of eTRS listing frequency, certification’s impact is roughly 50 percent greater than in markets at the seventy-fifth percentile. Second, we find that the eTRS badge has a greater effect in less concentrated markets, measured by category-level Herfindahl indices, consistent with the view that vertical differentiation is more valuable when product market competition is more intense. Certification’s impact is almost twice as large in eBay markets at the twenty-fifth percentile of seller concentration relative to those at the seventy-fifth percentile.

We then document the link between seller-level reputation and the value of certification. On eBay, a seller’s record of prior transactions is summarized as a publicly observed cumulative feedback score. Consistent with feedback serving as a substitute for eBay’s own quality certification, sellers with scores at the twenty-fifth percentile of seller feedback in the matched sample have an increase in sale probability from the badge that is 45 percent greater than its effect on sellers at the seventy-fifth percentile. Thus, certification’s impact is amplified for sellers who have not yet had the opportunity to demonstrate their reliability to prospective buyers via a record of satisfactory transactions.

Finally, we examine whether growth in eTRS certification generates a “business stealing” effect by reducing the sales of competitors. An increase in eTRS prevalence has a negative impact on other sellers in a category, particularly already certified ones, consistent with more widespread certification promoting greater competition among high-quality sellers.

Collectively, our findings highlight that the impact of quality certification on a particular seller—and hence the seller’s incentives to acquire and maintain certification—depends on characteristics of both the seller and the market, and generates a differential competitive impact on uncertified versus certified sellers. This also implies that quality assessment organizations should take market conditions into account in designing certification mechanisms, and that one-size-fits-all accreditations may not be optimal.

We contribute to a literature that has previously focused on the impact of quality certification in individual markets. Jin and Leslie (2003), for example, demonstrate that introducing restaurant hygiene report cards in Los Angeles resulted in consumers sorting toward cleaner establishments and a reduction in food-related illnesses. Jin and Leslie (2009) go on to compare certification's impact on chain versus non-chain establishments in a study of the interaction among certification, reputation, and cross-restaurant information spillovers, but they are unable to perform the type of cross-market analysis that forms the core of our paper.

In a similar spirit, Wimmer and Chezum (2003) find that certified race horses sell for higher prices and go on to have better racing careers than uncertified ones. Improved sorting can allow the benefits of certification to extend to low-quality products, as Tadelis and Zettelmeyer (2015) describe in a field experiment with used car auctions. Ramanarayanan and Snyder (2012) examine a public grading system for dialysis centers and show that centers with low grades serve fewer well-informed patients, and that low grades motivate centers to improve performance. There is evidence, however, that certification's impact may be attenuated by institution- or market-specific features. Xiao (2010) shows that certification has little impact on demand for childcare services if alternative quality information is already provided by firms, while Ho (2012) describes how inflation and inconsistency in grades may subvert the goals of restaurant hygiene systems. In a study of quality assurance mechanisms on eBay's platform in the United States, Hui et al. (2013) document the reduction in certification's impact when a new buyer protection program was introduced. Not all quality assurance programs are successful, however, as reported in Roberts' (2011) study of warranties for online tractor sales and Luca's (2011) research on Yelp restaurant ratings.¹ To our knowledge, we are the first to examine how differences in market-level attributes affect consumers' responses to certification.

More broadly, our paper contributes to a growing literature on the role of quality assurance mechanisms in stimulating trade. Much of this work has been motivated by the increasing importance of e-commerce and has been enabled by the ever greater ease with which detailed information may be collected from online marketplaces, most notably eBay (see, for example, Lewis 2011; Saeedi 2012; and Elfenbein, Fisman, and McManus 2012).

¹In some cases, firms' pursuit of quality certification can have unintended or perverse effects. Dranove et al. (2003) show that hospitals may decline to treat severely sick patients in order to avoid risking poor grades in New York's hospital "report card" program, while Forbes, Lederman, and Tombe (2015) explore the incentives of airlines to manipulate arrival times to improve records of on-time arrivals. In all of these studies (and in our own), the certifying organization is a government body or market-maker seeking to promote trade by providing information that the firms will not or cannot credibly provide independently. There is also a rich literature on self-interested third-party intermediaries (for example, Moody's) that provide quality information about firms, while also perhaps pursuing their own profit maximizing objectives or seeking to please the market participants (often firms) that underwrite their existence (e.g., Becker and Milbourn 2011).

I. Background and Setting

A. eBay's United Kingdom Auction Platform

Founded in the United States in 1995, eBay has emerged as one of the world's largest online marketplaces, with over 100 million users globally in 2012.² A United Kingdom site, www.ebay.co.uk, was launched in 1999 and in 2012 became eBay's second most active marketplace, after the US-based site. The main features of eBay's UK platform mimic those of the US-based platform, which have been described extensively in other studies (e.g., Bajari and Hortasçu 2004). Apart from minor country-specific differences (e.g., prices are listed in pounds), it is noteworthy that a larger fraction of items on the UK platform are listed for sale at fixed prices (instead of auctions) than on the US site³ and a greater proportion of items are listed by professional sellers.

B. The Top-Rated Seller (eTRS) Program

Buyers on eBay purchase products they cannot inspect from sellers with whom they cannot have face-to-face communication, and whom they trust to deliver the product after payment is received. This marketplace is thus vulnerable to seller misrepresentation of products, and service problems ex post (careless packaging, shipping delays, etc.), as well as outright fraud. Easy entry (and exit) from these marketplaces by sellers exacerbates these problems (Brown and Morgan 2006).

Since its founding, eBay has relied on a feedback system to allow buyers and sellers to generate public track records. When an eBay user is the seller in a transaction, his eBay feedback score increases by 1 if the transaction's buyer provides positive feedback; conversely, negative feedback reduces the score by 1. Buyers may receive positive, but not negative, feedback. Product listing pages include a display of the seller's feedback score and his fraction of positive feedback. A potential buyer also may visit a separate page to examine feedback ratings and comments the seller received on items sold in the previous 90 days.

Researchers have shown that this feedback system is vulnerable to manipulation (see, for example, Brown and Morgan 2006; Dini and Spagnolo 2009; Bolton, Greiner, and Ockenfels 2013). To augment it, eBay introduced Detailed Seller Ratings (DSRs) in 2007 and the eBay Top-Rated Seller (eTRS) program in 2009.⁴ Under the DSR program, buyers have the additional option of rating sellers following a transaction along four dimensions: Was the item received as described by the seller? Was the seller's communication effective? Was the product shipped in a timely manner? And were the shipping and handling charges reasonable? Each is

²These figures come from www.ebayinc.com/who (accessed 7/25/2012).

³For example, in the digital camera category on October 9, 2012, 98 percent of 9,634 items for sale in the United Kingdom were listed at a fixed price versus 80 percent of 55,241 items on the US site. For new digital cameras, the figures are 99.7 and 95.4 percent, respectively. See Einav et al. (2013) for a discussion of the decline of auctions on eBay.

⁴Prior to the eTRS program, eBay reported a seller's PowerSeller designation (none, bronze, silver, gold, platinum, or titanium) as a quality signal. PowerSeller levels were based largely on sales volume, however, and eBay designed eTRS certification to be more informative about per transaction expected quality.

rated on a five-point scale. Sellers' publicly viewable profile pages contain graphs that report the average scores along these dimensions as well as the number of past customers who provided ratings.

To become a Top-Rated Seller in the United Kingdom when the program started in late September 2009, sellers had to meet the following requirements: at least 100 transactions or £2,000 of sales in the prior year with UK or Irish buyers, a positive feedback rating of at least 98 percent, and minimum average DSR scores of 4.6 out of 5. Furthermore, Top-Rated Sellers were permitted to have no more than 0.5 percent or 2 instances of DSR ratings of 1 or 2 in the prior 3 months (if 400 or more transactions) or in the prior 12 months (if fewer than 400 transactions in the past 3 months). Finally, Top-Rated Sellers had to be registered as businesses on eBay, and had to include a comprehensive returns policy within each listing. Sellers who meet all of these certification requirements receive a badge to display on their profile and product listing pages; during the period we study, the eTRS badge was a small ribbon with the words "Top Rated Seller." One month into the program, the badge was also displayed next to the product title on the search results page (see Figure 1), enabling buyers to distinguish between eTRS and non-eTRS listings prior to viewing product listing pages. Beyond displaying the eTRS badge on the results and listings pages, eTRS sellers also received discounts (up to 20 percent) on fees paid to eBay, and as of late November 2009 improved search standing for listings in "Best Match" search results.⁵ For many users, Best Match was the default method through which search results were displayed, but users could also sort results by time, price, or distance. In our empirical analysis, we will thus take care to account for listings' visibility in search results.

While a seller's average DSR scores are publicly visible, eBay has an informational advantage relative to consumers in its ability to track instances in which the seller received very low DSR ratings on individual transactions. Since many sellers have high DSR averages and cumulative feedback scores, the eTRS badge conveys to buyers information on the (low) probability of a certified seller providing very poor service, which buyers could not infer themselves in the badge's absence. Correlation between DSR scores and eTRS status is positive but small (ranging from 0.12 to 0.15), and likewise for seller feedback.

Top-Rated Seller status is assessed on the twentieth day of each month, when eBay evaluates seller performance over the 3 months prior (or 12 months for sellers with fewer than 400 transactions). Changes in eTRS status are effective immediately, with open listings updated dynamically to reflect this. Thus, if a seller's performance on, say, May 3 pushes her performance metrics above the eTRS threshold, she will only receive a badge on May 20.⁶ Similarly, if a seller's slipping performance pushes her below the eTRS threshold, she keeps her badge until the next assessment. A seller's eTRS status applies across all categories in which she posts listings.

⁵Identical fee discounts were also provided to high-volume sellers whose performance was below the eTRS threshold.

⁶We encountered in the data only a small number of off-schedule adjustments to sellers' eTRS status, and we account for these changes in the same way we treat standard eTRS updates. Most transitions in and out of eTRS certification occur on the twentieth of the month.

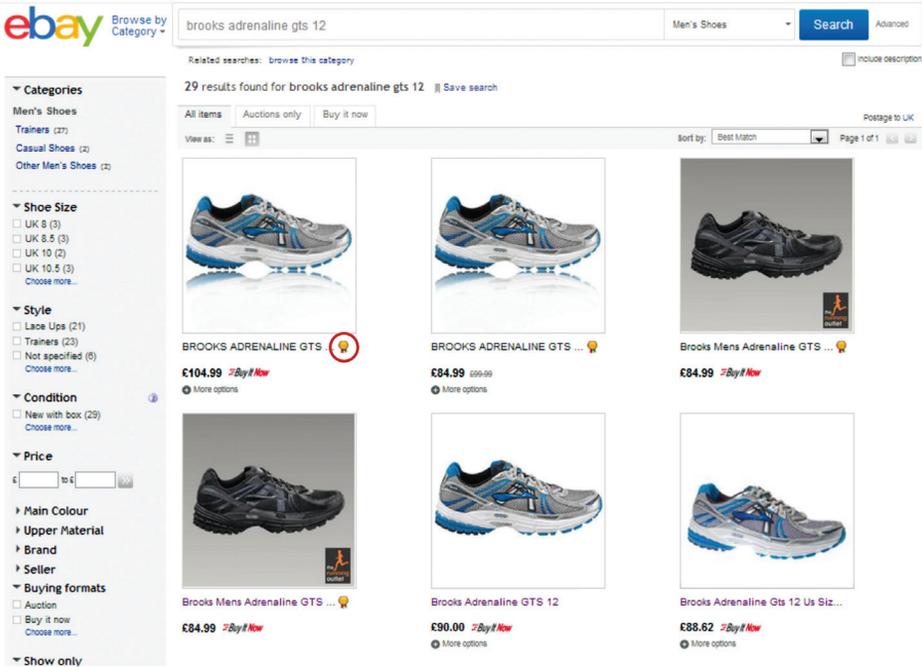


FIGURE 1. TOP-RATED SELLER ITEMS IN EBAY SEARCH RESULTS

Notes: The eTRS badge is displayed for four of the six products shown above. In the top left product we have added a circle around the badge.

II. Illustrative Model and Empirical Predictions

We begin with a theoretical framework to organize and interpret our analysis, focusing on intuitions gleaned from a more formal modeling exercise, included in the Appendix. As noted earlier, the relationships between market attributes and the impact of certification on demand are unclear a priori. This section's purpose is to show that the effects that dominate in our empirical analyses can be reconciled with standard models of competition augmented by unobserved quality.

We focus our attention on the effects of seller and market characteristics on consumer response to quality certification. For intensity of product market competition, we rely on prior art to frame our analysis. The first model we describe in the Appendix encapsulates the intuition that vertical differentiation is more valuable if firms are more similar in their horizontal attributes. Specifically, we consider a standard Salop (1979) circular city model where N evenly spaced firms sell to uniformly distributed consumers with binary demand. The true quality of seller i is captured by the parameter α_i , representing the likelihood that a seller offers an error-free transaction ($\alpha_i < 1$). Suppose, for simplicity, that seller types are either high (H) or low (L), with $\alpha_H > \alpha_L$. Absent feedback or certification, all sellers are assigned reliability $E(\alpha)$, reflecting the fraction of reliable sellers in the population. In this framework,

quality certification of seller i informs consumers that $\alpha_i = \alpha_H$. Holding prices fixed (as is the case within the groups of listings we consider), quality certification induces an increase in demand that is independent of initial market share, implying a larger proportional change to market shares when initial shares are smaller, as is the case when N is greater. Thus, *a seller experiences a larger percentage increase in sales due to certification when the seller is in a more competitive market.*

The second part of the Appendix, provides a discussion of the construction of $E(\alpha_i)$ based on information about sellers that accumulates over time. We presume consumers can observe public feedback, quality certification, and a measure of product market “riskiness;” this information is contained in Ω_i for seller i . We specify a two-period model to highlight an intuitive set of predictions that we analyze in the data. We retain the assumption that sellers differ in their unobserved quality—either high or low. Sellers operate in a single category, and the distribution of seller quality varies across categories (i.e., markets). This variability may result, for example, from product attributes such as ex ante verifiability of product quality and/or the likelihood of breakage in delivery. After each transaction, a buyer may provide publicly observable feedback. Prior to any transactions, eBay also may observe, with some probability, a perfectly informative signal of seller quality. We assume that, on the basis of this independent signal, a seller is given eTRS status if eBay learns that the seller is of high quality. Consumers use Bayes’ rule to form beliefs about $E(\alpha_i|\Omega_i)$ based on the feedback a seller has received combined with the seller’s eTRS status.

Our sparse model delivers two further predictions that relate seller and market attributes to the value of certification: (i) *certification’s impact decreases in the fraction of quality certified sellers in a market*, and (ii) *certification’s impact diminishes with the accumulation of buyer feedback, since the two serve as substitutes.* The former prediction is driven by the fact that since average seller quality is lower when certification is rare, buyers update more strongly in response to certification. Recall that eTRS requirements are the same regardless of the ease of selling in a particular category, so that differences in the category-level prevalence of certification are plausibly the result of differences in sellers’ unobserved ability to meet eBay’s eTRS standards in a particular market.

III. The Data

A. Sample Period and Data Extract

We examine a large data extract from eBay’s UK platform, which included data on individual listings that conclude between September 29, 2009 (the eTRS program’s first day) and October 31, 2010. We observe all listings in the sample period for 44,658 sellers who attain eTRS status at any point during the program’s first year. These “ever-eTRS sellers” account for 113 million listings. We also observe all listings of 1,982 additional sellers who approach but do not achieve eTRS status during the sample period (7 million listings). Finally, eBay provided an additional 10 percent sample of listings from the entire UK marketplace, containing 33 million listings from 2.2 million distinct sellers. We thus infer that the UK marketplace

hosted nearly 450 million listings during the sample period, with 25 percent coming from sellers who held eTRS status at some point during this time.

The listing data provide information on a product and its seller; each listing's selling format (e.g., true auction or fixed price); the number of units available and sold; listing details such as start date, end date, number of photos displayed, and shipping fees; the fixed price or auction starting price; and for auctions the data provide the number of bids, selling price, and maximum bid value. We also observe the number of times a listing was shown to consumers (an "impression") as part of a list of search results, and the number of times consumers click-through to the listing's web page (a "view"). Seller and listing characteristics, including the presence of an eTRS badge, affect eBay's algorithm for serving search results to consumers, so it is crucial to control for the numbers of impressions and views to separate the informational effect of the badge from its effect on visibility in searches. For the sellers with complete listing data, we observe their complete eTRS history. We see detailed seller ratings (DSRs) and feedback scores monthly. Finally, we observe the annual and quarterly summaries of transactions and revenue that eBay uses to evaluate a seller's eTRS status.

We observe each item's location in eBay's hierarchy of product categories, and identify a product's market based on its "leaf category," the most specific classification in the eBay hierarchy. Our listings are drawn from over 8,000 leaf categories (e.g., 4GB Apple iPod Mini), which themselves are members in 1 of 33 top-level categories (e.g., Consumer Electronics). The variety within a leaf category is determined in part by the eBay market thickness for a class of products, and there will be variation across categories in the substitutability of products that are grouped together.

To summarize weekly activity on the UK site, we construct several market-level panels using the full listing data, the seller characteristics, and the leaf category codes. Within each week and leaf category, we count the listing days associated with each active seller, the number of units sold, revenue collected, and the seller's badge status. We then aggregate these data within a market-week, and calculate market concentration measures like Herfindahl-Hirschman Indices (HHIs) and shares of all listings, sales, and revenue that originate with eTRS-badged sellers.⁷ We assume that a category's market structure and eTRS share are uncorrelated with unobserved factors that might drive category-level differences in consumers' responses to the badge.

B. Matching Procedure and Sample Characteristics

From the data extract we assemble two types of matched listings, labeled "ST" (seller-title) and "STP" (seller-title-price) matches. An ST-match consists of a group of two or more listings from a single seller that uses the same title, subtitle, and

⁷ We combine the 10 percent and 100 percent samples of sellers' listings by weighting the latter group's listings and transactions by 0.1 in calculating market activity summaries, category HHI measures, and share of eTRS listings within a category.

selling format (e.g., fixed price).⁸ An STP-match is a group of items listed by the same seller that use the same title, subtitle, selling format, *and* posted price or start price, as appropriate based on selling format. We have used this approach in prior research (Elfenbein, Fisman, and McManus 2012), as have Einav et al. (2015). Of the 113 million listings from ever-eTRS sellers in our data extract, over 100 million can be matched in this way. In many matches, a seller's eTRS status differs among listings as a result of an individual seller having gained or lost eTRS status. Under the assumption that this within-match variation in eTRS status is exogenous with respect to demand, product, and seller characteristics, these differences provide an opportunity to credibly estimate certification's impact on demand. This is plausible in our setting as small changes in seller attributes around the eligibility threshold produce a discrete change in eTRS status. (One potential concern is that STP-matches might not represent clean certification "experiments," as sellers can adjust prices in response to receiving the badge. Empirically, we find no evidence that sellers respond in this way, as we discuss in Section IV below.)

Of the full collection of 100 million matched listings, we exclude 38 million auction-style listings from our analysis, since they account for considerably less trade on eBay UK than fixed price offerings.⁹ We drop 18 million fixed-price matches with an average posted price below £4.95 or above £500. About 17 percent of the remaining listings are eliminated because the seller did not change eTRS status during the sample period (most often because the seller received certification immediately and never lost it). We also drop a small number of observations due to missing data, unusually large quantities offered within single listings, or other irregularities. Our final sample includes approximately 31 million listings, and we further restrict our attention to the 16.3 million listings that have within-match variation in the quantity of items sold (i.e., whether an item sold in the case of single-item listings).¹⁰

In Table 1, panel A, we summarize the listing-level characteristics of the final set of ST matches. Fifty-two percent have a badge when the listing ends and 27 percent finish with the sale of one or more units. Many fixed-price listings feature multiple units, in which case our data provide the quantities available and sold. At the data medians, listings are active for 10 days, include one photo of the item for sale, and have a shipping fee of £1.50. Listings typically appear in categories that are fairly competitive based on HHI, though there is wide variation across categories. In our sample, the average category-level eTRS share is 24 percent. The mean price

⁸ eBay identifies three major selling formats: "fixed price," "store fixed price," and "auction." In our discussion we use the term "fixed price" to describe the first two selling formats, although we separate these formats in creating the matches. This implies that format-specific differences in average listing outcomes are picked up by the match fixed effects we introduce below.

⁹ The fixed price listings captured by our matching procedure result in transactions that have total revenue value that is six times as large as the value of the matched auction listings.

¹⁰ Groups of listings with identical values in the outcome variable are not informative on the effect of certification, yet the estimated coefficient on eTRS status would be lower in our linear probability model if we retained these observations. The intuition is as follows: a binary-outcome econometric model's latent equation can contain a nonzero badge effect while allowing a large negative fixed effect to match an empirical result that all items in a group fail to sell. The linear probability specification makes no distinction between the latent and observed models, however, and when no objects in a group sell, the least-squares estimates of all coefficients are pushed toward zero. (The same argument applies to a group of listings in which all items sell.)

TABLE 1—SUMMARY STATISTICS FOR MATCHED SAMPLE

	Mean	Median	SD	Min.	Max.
<i>Panel A. Listings (Observations = 16.3M)</i>					
Badge ($Y = 1$)	0.52	1	0.5	0	1
Sale ($Y = 1$)	0.27	0	0.45	0	1
Quantity available	21.82	4	81.27	1	1,000
Quantity sold	0.67	0	3.41	0	1,000
Offered price	29.1	13.99	48.25	0	3,499
Price Sold (<i>Observations = 4.5M</i>)	26.32	12.99	43.01	0.99	1,117.93
Number of impressions	4,442.78	1,494	9,778.59	0	1.7E+6
Number of views	41.48	10	132.2	0	26,029
Shipping fee	2.59	1.5	5.46	0	6,000
Number of photos	1.16	1	0.89	0	12
Scheduled length	17.78	10	11.76	1	30
Actual length	15.38	10	11.51	0	47
Category HHI by listings	575.68	267.76	950.77	1.74	10,000
Category HHI by quantity	843.39	384.72	1,298.68	0	10,000
Category eTRS share	0.24	0.21	0.15	0	1
<i>Panel B. Sellers (Observations = 22,801)</i>					
Number of listings	1,364.55	134	10,774.18	2	643,306
Type-ST matches	95.71	13	1,148.84	1	99,290
Type-STP matches	168.80	17	3,004.44	1	264,949
Feedback score	4,267.67	1,562.02	11,204.26	4.64	737,378
Transactions in last 3 months	538.18	202.11	1,456.88	0	87,360
Revenue in last 3 months	9,845.55	3,472.40	45,989.27	0	4,851,372
Category listing share	0.03	0.01	0.07	1.31E-06	0.98
Category sold share	0.04	0.02	0.08	4.9E-06	0.98

Notes: Panel A contains summary statistics on ST matches in which there is variation in quantity sold. Panel B includes data from all sellers who contribute to the ST matches; some of these sellers' listings do not vary in quantity sold.

is £26.32 (median £12.99) for successful transactions, slightly below the average posted price among the matched listings. The average number of units sold per listing (0.67) is greater than the overall success rate due to successful sales of multiple units from a single listing.

Our matches come from 22,801 sellers, whose characteristics are also summarized in Table 1, panel B. The median seller has 134 listings across 13 ST matches. The mean and median feedback scores are 4,268 and 1,562, respectively, while the median seller has 202 successful transactions per quarter and £3,472 in revenue; the means for these variables are about three times the median values. Finally, despite their size, sellers generally have low market shares within each item's category.

IV. Empirical Analysis

We first demonstrate that eTRS status improves seller outcomes, then examine three predictions that relate certification's value to market structure and seller reputation: (i) eTRS status will have a greater effect on sales probability in product categories where certification is rare; (ii) eTRS status will be more valuable in more competitive categories; and (iii) the positive effect of certification will be amplified for sellers with limited transaction histories. We focus throughout on the matched data where title, subtitle, and seller are identical (in most cases price as well) in order to reduce concerns of unobserved differences across sellers or products.

The basic econometric specification for the matched analysis is

$$(1) \quad \begin{aligned} \text{Sale}_i = & \mu_m + \beta \mathbf{X}_i + \gamma \text{Badge}_i + \delta \mathbf{Market}_i \times \text{Badge}_i \\ & + \theta \mathbf{Seller}_i \times \text{Badge}_i + \varepsilon_i. \end{aligned}$$

The dependent variable *Sale* is an indicator for whether a listing *i* ends with a sale. In some specifications we replace *Sale* with transformations of *Quantity*, the count of items sold in a listing. There is a fixed effect μ for each group (*m*) of matched items (which effectively absorbs both product and seller fixed effects); the vector \mathbf{X} captures additional observable variation in listing characteristics, such as impressions, number of photos, and listing duration. *Badge* captures a seller's eTRS status (*Badge* = 1 if a seller is certified, 0 otherwise), with the parameter γ as the marginal effect of the badge on *Sale*. When we examine the heterogeneous effects of certification, we also include the interactions between *Badge* and vectors of market and seller characteristics. The error term ε accounts for additional variation in outcomes across listings. In estimating the parameters in (1), we cluster standard errors at the seller level.

We estimate (1) using standard linear regression methods. While this approach does not account for the discrete nature of the dependent variable *Sale*, we are able to include a large number of match-level fixed effects (μ) that would be computationally demanding in nonlinear models. Further, this approach sidesteps the problems associated with interpreting interaction terms in nonlinear models. Where possible, we focus on STP matches to eliminate concerns about estimating γ when sellers might reprice following a transition in certified status. Potential concerns about repricing—and hence whether price variation is exogenous in ST matches—are alleviated by additional empirical analysis, reported in Table S1 in the Supplementary Online Appendix,¹¹ which shows that a change in badge status results in a precisely measured zero average change in posted price (though sellers might change their behavior through the number and value of items listed on eBay rather than price-setting after gaining eTRS status).

A. Base Effects of the eTRS Badge

Table 2 shows the results from our initial estimation of (1), restricting the interaction coefficients δ and θ to equal zero. We include as controls a listing's timing, duration, shipping fees, and number of photos, but for brevity exclude these parameters from Table 2. See Table S2 in the Online Appendix for the full set of parameter estimates.

Specifications 1–4 in Table 2 employ *Sale* as the dependent variable, while specifications 5–8 use two *Quantity*-based measures, $\log(1 + \text{Quantity})$ and $\text{Quantity}/(\text{Quantity Available})$. In specification 1, we estimate the full effect of *Badge* on sale probability in Seller-Title-Price (STP) matches, inclusive of the

¹¹www.unc.edu/~mcmanusb/AEJMicro-EFM-SuppApp.pdf.

TABLE 2—BASE RESULTS FROM MATCH ANALYSIS

Dependent variable Specification	Sale			
	(1)	(2)	(3)	(4)
<i>Badge</i>	0.0219*** (0.00166)	0.0182*** (0.00164)	0.0193*** (0.00140)	0.0264*** (0.00505)
$\log(\textit{Price})$				-0.312*** (0.0200)
$\log(\textit{Price}) \times \textit{Badge}$				-0.00220 (0.00148)
$\log(\textit{Impressions})$		0.0404*** (0.00280)	-0.0253*** (0.00401)	0.0374*** (0.00284)
$\log(\textit{Views})$			0.114*** (0.00333)	
Match type	STP	STP	STP	ST
Observations	14,070,450	14,070,450	14,070,450	16,281,095
R^2	0.008	0.016	0.034	0.025
Number of matches	1,597,977	1,597,977	1,597,977	1,656,439
Dependent variable Specification	$\log(1+Q \textit{ sold})$		$Q \textit{ sold}/\textit{available}$	
	(5)	(6)	(7)	(8)
<i>Badge</i>	0.0171*** (0.00156)	0.0313*** (0.00517)	0.0119*** (0.00148)	0.0157*** (0.00522)
$\log(\textit{Price})$		-0.294*** (0.0195)		-0.244*** (0.0198)
$\log(\textit{Price}) \times \textit{Badge}$		-0.00455*** (0.00158)		-0.000734 (0.00148)
$\log(\textit{Impressions})$	0.0584*** (0.00241)	0.0569*** (0.00250)	0.00703*** (0.00136)	0.00542*** (0.00152)
$\log(\textit{Views})$				
Match type	STP	ST	STP	ST
Observations	14,070,450	16,281,095	14,070,450	16,281,095
R^2	0.046	0.054	0.014	0.027
Number of matches	1,597,977	1,656,439	1,597,977	1,656,439

Notes: Robust standard errors, clustered by seller, are in parentheses. *Badge* is an indicator variable denoting whether the seller had eTRS status at the time a listing closes; *Price* is the listing's posted price; *Impressions* is the number of times a listing was shown to buyers as part of a search result; and *Views* is the number of times a listing received a "click-through" and was viewed by a potential buyer. In addition to the listed explanatory variables, the specifications also include the following controls for listing characteristics: the logarithm of shipping fee, whether the listing occurred in the first month of the eTRS program, a quartic time trend, and sets of indicator variables to control flexibly for each of the listing's duration, its ending day-of-week, and its number of photos. See online Appendix Table S2 for the full set of coefficient estimates for specifications 2, 5, and 7. STP match denotes sets of listings from a seller with identical titles, subtitles, and posted prices, while ST matches denotes listing sets from a seller with identical titles and subtitles.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

effect on search ordering. Badged listings are 2.2 percentage points more likely to sell, 8 percent higher than the base success rate of 27 percent for unbadged items. Specification 2 uses the log of one plus the number of impressions to control for the effect of the badge on a listing's position in search results, and we interpret the new

estimate of γ to be the “informational” effect of eTRS status. We still find a relatively large impact (7 percent) of the badge on sale probability.¹² The estimate of γ is virtually unchanged when we include the log of views (specification 3), although a consumer’s decision to view an item could be affected by the quality-relevant informational value of a badge. In specification 4, we analyze ST matches, which allow for variation in the fixed price of the product, and also includes the interaction of $\log(\textit{Price})$ and *Badge* for price effects to vary with badge status. We find that the sale probability falls significantly with the log of an item’s fixed price, while the coefficient on the interaction term is small in magnitude and statistically insignificant, indicating that both badged and unbadged items have similar price elasticities.¹³

In specifications 5–8, we report analogous specifications using two quantity-based measures as the dependent variable, $\log(1 + \textit{Quantity})$ and $\textit{Quantity}/(\textit{Quantity Available})$, and obtain results that are qualitatively similar to those on sale probability. The coefficient estimates of specifications 5 and 7 imply that certified listings have values of the outcome variable that are 4 percent and 8 percent greater, respectively, than items with *Badge* = 0.

In Table 3, we provide further robustness checks for our estimates of γ . To account for possible within-match heterogeneity in listing attributes, we repeat our main analyses with a much more stringent match requirement, where within a match all listings are identical in terms of title, subtitle, and price, number of photos, shipping fee, scheduled length, and quantity available, and we report these results in column 1. This has little effect on our estimate of certification’s impact on demand. In column 2, we limit our sample to listings with a change in eTRS status between the current month and either the preceding or next one, and in column 3, we include only listings that conclude in the ten days prior to an eTRS status change and the ten days following it. By focusing on the window immediately around status changes, we are better able to isolate demand-side responses to *Badge* from possible post-certification seller responses or time trends. We obtain point estimates that are very similar to those in Table 2, and when we apply the appropriate baseline sale probabilities (based on the sample window) to these coefficients, we find that the eTRS badge shifts the sale probability by 8 percent in both cases. Finally, in column 4, we use the fraction of time during an item’s listing period when the seller held eTRS status (rather than an indicator variable *Badge* at the end of listing); once again the point estimate indicates a similar impact of eTRS status.

Table S3 in the online Appendix shows the badge’s effects on impressions and views—eTRS status has a positive effect on both outcomes. Once impressions are

¹²To further ensure that the coefficient in this specification can be attributed to the informational content of eTRS, we have tried specifications that control more flexibly for impressions. When we replace $\log(1 + \textit{impressions})$ with nine dummy variables that capture the deciles of percentage differences between a listing’s number of impressions and the match’s mean impressions, we obtain a nearly identical estimate of γ .

¹³As noted above, we take this price variation to be exogenous, possibly as a result of experimentation by the seller. (See Einav et al. 2015 for an extensive discussion of the matched listing approach as applied to eBay’s US platform, and its usefulness in estimating the slope of demand.) The presence of product-specific fixed effects and time trend controls allay many of the usual concerns about price endogeneity. Our finding of zero average price response from receiving the badge, noted above, provides some further support in the use of this variation in our demand-side analysis.

TABLE 3—ROBUSTNESS OF BADGE EFFECT

Dependent variable	Sale			
	(1)	(2)	(3)	(4)
<i>Badge</i>	0.0210*** (0.00188)	0.0186*** (0.00129)	0.0275*** (0.00281)	
Share of listing time with <i>Badge</i> = 1				0.0264*** (0.00198)
$\log(\text{Impressions})$	0.0439*** (0.00350)	0.0373*** (0.00236)	0.0490*** (0.00521)	0.0402*** (0.00281)
Match type	STP + Listing characteristics	STP + Recent month eTRS change	STP + 20-day window eTRS change	STP
Observations	11,939,804	4,451,912	551,446	14,070,450
R^2	0.010	0.014	0.019	0.016
Number of matches	1,494,211	678,305	122,932	1,597,977

Notes: Robust standard errors, clustered by seller, are in parentheses. Coefficients on additional control variables are omitted to conserve space. In addition to the listed explanatory variables, the specifications include the control variables discussed in the notes to Table 2 and provided in online Appendix Table S2.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

accounted for, *Badge* has no incremental impact on listing views. These regressions, along with the 20 percent difference in estimates of γ between specifications 1 and 2 in Table 2, demonstrate that controlling for the prominence effects of certification is critical to assessing its informational impact.

We may also calculate consumers' tradeoff between certification and price using our estimates of the badge effect and the slope of demand. From the parameter estimates of specification 4 in Table 2, we may calculate the predicted sale probability for each unbadged item in our sample. We then compute the value of Δ such that an unbadged item with price p_u has the same sale probability when offered by an eTRS seller for $p_u + \Delta$. For the fixed-price listings in the matched sample, the mean and median of Δ are £1.65 (6.7 percent) and £0.89 (6.8 percent), respectively. Computing the same price increments using the estimates from specification 6 to equate the expected quantity sold in a listing, we find the mean and median of Δ are £1.27 and £0.88, respectively.

We may then calculate sellers' value from eTRS status. Using estimates from specification 5 in Table 2, we calculate incremental revenue for listings in the STP sample due to the eTRS badge. There are 7.4 million eTRS listings, with an average sale price of £26.29, and the coefficient estimate of 0.0171 implies an increase in quantity sold of 0.0285 per listing. Hence, the increment to revenue is £5.6 million (£7.4 million if we remove controls for impressions). Since the STP sample represents about 30 percent of all fixed price listings, the total incremental revenue produced by the badge is likely to have been significantly larger.¹⁴

¹⁴In the year following the introduction of the program, a total of 25.3 million fixed-price items were listed with eTRS status, and the average sale price for listings by sellers who received the badge at some time during this period was £27.92.

TABLE 4—INTERACTION OF CERTIFICATION WITH MARKET AND SELLER CHARACTERISTICS

Dependent variable	Sale					
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Badge</i>	0.020*** (0.0016)	0.0093*** (0.0023)	0.031*** (0.0029)	0.048*** (0.0086)	0.042*** (0.0088)	
<i>Listing HHI</i> × <i>Badge</i>	−0.0037*** (0.0011)					
log(<i>Listing HHI</i>) × <i>Badge</i>		−0.0067*** (0.00099)			−0.0046*** (0.00097)	−0.0063*** (0.0010)
<i>eTRS Share</i> × <i>Badge</i>			−0.048*** (0.0088)		−0.034*** (0.0094)	−0.0298*** (0.0088)
log(<i>Feedback</i>) × <i>Badge</i>				−0.0034*** (0.0011)	−0.0024** (0.0011)	−0.0024** (0.0011)
Includes <i>Badge</i> × metacategory fixed effects?	No	No	No	No	No	Yes
Observations	14,070,450	14,070,450	14,070,450	14,069,456	14,069,456	14,069,456
R^2	0.016	0.016	0.016	0.016	0.016	0.016
Number of STP matches	1,597,977	1,597,977	1,597,977	1,597,653	1,597,653	1,597,653

Notes: Robust standard errors, clustered by seller, are in parentheses. *Badge* denotes whether the seller had eTRS status at the time a listing closes, and *Listing HHI* denotes week × leaf category Herfindahl-Hirschman Indices of concentration calculated using sellers' shares of listings. *HHI* is expressed on a 0–10 scale to facilitate interpretation of the coefficients. *eTRS Share* is the percentage of listings in a week × leaf category that are offered by sellers who have the eTRS badge at the time the listing closes. *Feedback* is a seller's eBay feedback score. Coefficients on additional control variables are omitted to conserve space. In addition to the listed explanatory variables, the specifications include the control variables discussed in the notes of Table 2 and provided in online Appendix Table S2.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

B. Market Characteristics, Seller Characteristics, and the Impact of eTRS Certification

Our main contribution is analyzing how product market structure and seller characteristics affect certification's impact on demand. We thus turn our attention to (1), focusing now on the interaction between *Badge* and market/seller characteristics (parameters δ and θ). The model sketched in Section II emphasizes that the impact of certification on demand should decline in market concentration, the proportion of certified sellers in a category, and a seller's feedback history.

Table 4 presents our main results, and we follow it with a description of a number of alternative specifications and robustness checks that are reported in the online Appendix. Throughout, we control for impressions, so we interpret our estimates of the market and seller characteristics' effect on badge value as representing changes its informational value.

Column 1 includes the interaction of *Badge* and the category-level variable *Listings HHI*, which captures the impact of industry concentration on the badge's value. We divide all HHI statistics by 1,000 for ease of reading. The coefficient on the badge-HHI interaction is -0.0037 , significant at the 1 percent level; this attenuating effect of concentration is consistent with the intuition captured by the circular city model outlined in Section II. The coefficient implies that the badge's effect

at the twenty-fifth percentile of *Listings HHI* (118) is 10 percent greater than the effect at the seventy-fifth percentile *HHI* value (610). *Listings HHI* has an extremely long right tail; we therefore employ the transformed variable, $\log(\textit{Listings HHI})$ (which has a smooth, normal distribution) in column 2; these results imply a much larger reduction in the badge's impact (47 percent) between the twenty-fifth and seventy-fifth percentiles of *HHI*.

To examine how the badge's value varies with the category-level eTRS prevalence, we include in column 3 the interaction of *Badge* and *eTRS Share*, the fraction of listings (in the most narrowly defined category) with eTRS status. The coefficient of -0.048 (significant at the 1 percent level) implies that moving from the twenty-fifth to seventy-fifth percentile of *eTRS Share* (14 percent to 31 percent) reduces the badge's effect by a third. This result is consistent with the intuition outlined in Section II, where consumers infer that categories with a higher rate of certified listings have fewer quality problems, thus reducing the difference in expected performance between badged and unbadged sellers.

Column 4 examines whether certification and seller feedback are substitutes, by including $\log(\textit{Feedback}) \times \textit{Badge}$. Consistent with this prediction, the interaction term's coefficient is negative, taking a value of -0.0034 (significant at the 1 percent level). Moving from the twenty-fifth to seventy-fifth percentile of seller feedback implies a 30 percent reduction in the badge's impact. Column 5 includes all three interactions simultaneously, where we find that all retain significant explanatory power.¹⁵

To address the concern of category-level attributes correlated with eTRS share or concentration that also affect eTRS value, column 6 includes fixed effects for each of eBay's 33 meta-categories interacted with *Badge*. These interactions absorb, at the meta-category level, any unobserved category differences correlated with concentration or eTRS prevalence. Our findings are largely unaffected by these additional controls.

In each specification above, our estimate of certification's impact on demand is based on the assumption of a linear relationship between the badge and market or seller attributes. We allow for greater flexibility in these relationships in Table S4 of the online Appendix, which also provides a more straightforward interpretation of effect magnitudes. When we include interactions of *Badge* and indicator variables for *HHI* quartiles, the coefficient estimates imply an effect of certification that is more than twice as large for listings in the lowest *HHI* quartile relative to the highest. When we interact *Badge* and indicator variables for *eTRS Share* quartiles, the estimated coefficients monotonically decrease in *eTRS Share*, with an effect that is 2.5 times higher for low *eTRS Share* versus high *eTRS Share* categories. Finally, the interaction of *Badge* with *Feedback* quartile dummies generate coefficients that are decreasing with seller feedback, and the implied effect of certification is over twice as large for low- versus high-feedback sellers. In Table S5 we repeat

¹⁵If we use *Listings HHI* instead of $\log(\textit{Listings HHI})$, its interaction with *Badge* loses its significance when $\textit{eTRS Share} \times \textit{Badge}$ is included. One concern with this specification is that very highly concentrated categories will, almost by definition, have extreme values of *eTRS Share*, mechanically increasing variation in this variable due to market structure rather than category risk. This occurs most strongly at extreme values of *HHI*, which are less influential when we use the log specification.

our main analyses but with alternative measures of market concentration based on quantity sold. Our findings are virtually unchanged when using these quantity-based measures.

C. Competitive Impact of eTRS Badge

The positive impact of certification on demand that we have documented could represent purchases that otherwise would not have occurred, or sales that come at the expense of competitors. Moreover, if certification's effect on demand comes at competitors' expense, some sellers may suffer more than others when rivals receive certification. For example, if eTRS sellers gain market share primarily from unbadged rivals, then certification may bolster market concentration among (differentiated) high-quality firms. Alternatively, if sellers who gain the badge draw market share only from badged rivals, this may promote robust competition at the high-quality end of the market. A decisive analysis of the full welfare effects of such shifts in market share would require a more detailed model and appropriate data, but we present suggestive patterns on these issues that may indicate useful directions for future work.

We examine the "business stealing" effect of certification using a variant on equation (1) that examines sales probability within a given STP match as category-level eTRS prevalence varies, *holding seller eTRS status constant*. We identify the impact of changes in eTRS prevalence based on category-level changes in the prevalence of eTRS across a set of STP matches. For example, a seller may post one listing that closes on the fifteenth of the month (before eTRS update) and a second identical one closing on the twenty-second. As a result of eTRS updates on the twentieth, eTRS prevalence changes within the listing's category, allowing us to identify the effects of eTRS prevalence using within-match estimates. Our estimating equation is

$$(2) \quad \text{Sale}_i = \mu_m + \mathbf{X}_i\beta + \varphi e\text{TRS Share}_i + \varepsilon_i,$$

where Badge_i is held constant within m . We estimate (2) with an expanded dataset, including matched listings from the 10 percent platform-wide sample and those of eTRS sellers whose badge status is unchanged. We retain the same data filters as for our main matched sample (i.e., fixed price listings only, price between £4.95 and £500, and so forth).

We present our estimation results in Table 5. In specification 1, we report the estimate of φ in equation (2) within the set of STP matches where $\text{Badge} = 0$. The coefficient is negative and significant at the 1 percent level, indicating that a 10 percentage point increase in category-level prevalence of eTRS listings reduces the probability of sale for noncertified sellers by 0.4 percentage points. (This change in eTRS share is approximately equal to the within-category standard deviation of monthly eTRS shares, 0.087.) In specification 2, which includes STP matches where $\text{Badge} = 1$, the coefficient on eTRS prevalence is nearly twice as large, indicating a greater business stealing effect from certified sellers. The relative sizes of φ across unbadged and badged sellers may be due to a crowding at the top end of the quality ladder when the eTRS share increases.

TABLE 5—IMPACT OF COMPETITORS' eTRS CERTIFICATION ON SALE

Fixed seller eTRS status Specification	Badge = 0 (1)		Badge = 1 (2)	
<i>eTRS Share</i>	-0.0437*** (0.00852)		-0.0702*** (0.0114)	
Observations	6,735,047		9,071,017	
R^2	0.018		0.013	
Number of STP- <i>Badge</i> matches	894,593		1,171,894	
Seller characteristic Specification	Small seller (3)	Large seller (4)	Small seller (5)	Large seller (6)
<i>eTRS Share</i>	-0.0530*** (0.0120)	-0.0409*** (0.0136)	-0.0828*** (0.0131)	-0.0559*** (0.0147)
Observations	2,494,747	3,269,265	3,050,929	6,020,088
R^2	0.016	0.022	0.012	0.015
Number of STP- <i>Badge</i> matches	400,839	407,674	495,566	676,328
Product characteristic Specification	Low price (7)	High price (8)	Low price (9)	High price (10)
<i>eTRS Share</i>	-0.0377*** (0.0105)	-0.0503*** (0.0108)	-0.0637*** (0.0157)	-0.0773*** (0.0126)
Observations	3,561,709	3,173,338	4,575,252	4,495,765
R^2	0.021	0.015	0.014	0.012
Number of STP- <i>Badge</i> matches	451,182	443,411	604,103	567,791

Notes: Robust standard errors, clustered by seller, are in parentheses. Specifications 1, 3, 4, 7, and 8 hold fixed *Badge* = 0 within a match. The other specifications have *Badge* = 1 for all listings within a match. *eTRS Share* is the percentage of listings in a week \times leaf category that are offered by sellers who have the eTRS badge at the time the listing closes. Seller size models (specifications 3–6) split the seller population based on quarterly revenue. Low/High Price models (specifications 7–10) divide the matches by whether the mean price within-match is above or below the sample median. Coefficients on additional control variables are omitted to conserve space. In addition to the listed explanatory variables, the specifications include the control variables discussed in the notes of Table 2 and provided in online Appendix Table S2.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

We next disaggregate sellers by sales volume in the month prior to listing. In specifications 3–4, we compare the effect of eTRS prevalence on unbadged sellers, disaggregated based on whether their revenue for the previous three months was above £3000, and in specifications 5–6, we present analogous results for badged sellers. In both pairs of results, the coefficient on eTRS prevalence is more negative for smaller sellers, indicating an attenuated effect of business stealing on high-volume sellers. In terms of its percentage impact on sales probabilities, this difference is smaller once we account for the baseline difference in sales probability (0.33 versus 0.28 for small versus large sellers, respectively).¹⁶ Overall, these results suggest that larger sellers may be more successful at differentiating themselves from their badged competitors. In specifications 7–10, we examine whether the effect of eTRS prevalence

¹⁶The listings included in the analysis in Table 5 must have variation in *Sale* within a fixed badge status, which yields the differences between these success rates and the STP-sample value of 0.27 reported on Table 1.

differs for high- versus low-priced items, subdividing the sample based on whether a listing's price is above £14. For both badged and unbadged sellers, the coefficient is more negative for high-priced listings (the sales completion probability is 0.29 for both listing types). This greater business-stealing effect of certification among high-priced listings could result from a greater concern over seller misrepresentation or opportunism.

V. Conclusion

As Arrow (1972) famously noted, "virtually every commercial transaction has within it an element of trust." Private-party certification is one mechanism that enables transaction partners to overcome this trust problem. In theory, certification improves market performance by providing information that enables buyers to assess seller quality and attach differences in willingness-to-pay to sellers of various quality levels.

Our results raise a number of interesting questions about the impact of certification programs on the evolution of markets. The differential benefit of certification for new versus established players suggests that certification may enable high-quality entrants to grow faster, making concentrated markets more competitive. Alternatively, new low-quality entrants (sellers that will not attain certification) are at a greater disadvantage in the presence of many other certified sellers, potentially leading them to exit rapidly or deterring their entry altogether. Our results also suggest that these dynamics will be affected by the design of the certification program, specifically whether the quality threshold is set so as to enable few versus many market participants to obtain it. The dynamic relationship between firm reputation and size, market concentration, and certification design may be a fruitful area for future research.

Our findings have implications for the impact of quality certification within and across many markets. We predict, for example, that Consumer Reports recommendations or a Good Housekeeping Seal of Approval have a different effect on sales of irons versus espresso makers due to differences in market attributes. Moreover, our results suggest that the value of accreditations for high-stakes decisions like hospitals or childcare may vary with local market structure conditions, which could expose some shortcomings in one-size-fits-all public policy recommendations. Our full set of findings is applicable in many online markets where customer feedback and expert assessments are commonplace. The technology review site CNET, for example, offers links to all products in a given category (e.g., desktop computers), while also highlighting those that its reviewers have highlighted as "Best in Category." As with eBay, we expect the value of CNET's certification would vary by the extent and quality of reader feedback, and also a given category's market structure. By contrast, Yelp, whose paid business subscribers are given the opportunity to choose which reviews to feature prominently, may find that the value of affecting the presentation of quality assessments changes as more competing businesses subscribe to the service.

Our work may be extended in a number of directions. In this paper, we have focused on a demand-side response to certification. But it is possible—indeed likely

—that over time eTRS status may induce sellers to enter new categories or increase the number of listings they post. We defer this supply-side analysis to future work.

Finally, while our emphasis in this paper has been on the heterogeneous effects of certification as a function of market conditions, one could similarly analyze how the benefits of other quality signals vary with industry conditions. Empirically, our setting has the advantage of discrete shifts in certification status that allow us to generate credible estimates of certification's benefits. In future work we hope to develop approaches to extend our work to assess the heterogeneous impact of firm reputation and other quality assurance mechanisms.

APPENDIX: THEORY

I. Product Market Competition and the Value of Certification

We model product market competition as a variant on Salop's (1979) circular city. Consumers are distributed with unitary mass around a circle of unit circumference. N sellers, each with marginal cost of c , are spaced evenly around the circle. For consumers, we assume linear travel costs t and utility from the good of u . Sellers differ in their ability to successfully complete transactions. We assume that α_i represents the probability that the i_{th} seller completes transactions successfully. We model reliability as being a binary attribute, with low-type (high-type) sellers having reliability of α_L (α_H). The reliability parameters satisfy $0 < \alpha_L < \alpha_H < 1$. In the absence of additional information on individual seller quality, consumers expect utility from the good of $E(\alpha)u$, where the expectation of α simply comes from the fraction of high types in the seller population, which we assume to be ϕ . Without additional information on a seller, $E(\alpha) = (1 - \phi)\alpha_L + \phi\alpha_H$.

We assume travel costs are sufficiently high that all consumers buy from one of the two closest firms. A consumer located a distance x from seller i is indifferent between buying from i and his neighbor $i + 1$ if

$$E(\alpha_i)u - p_i - tx = E(\alpha_{i+1})u - p_{i+1} - t(1/N - x).$$

For the simple symmetric equilibrium case where all firms have the same expected reliability, $E(\alpha)$, seller i 's demand is given by $q_i = (p - p_i + t/N)/t$, where p is the equilibrium price. Maximizing profits with respect to p_i generates $p = c + t/N$, with market shares of $1/N$.

Within the context of this symmetric equilibrium, consider the effects of seller i obtaining certification from eBay, assuring consumers that $\alpha_i = \alpha_H$. Holding the prices of other sellers constant, i 's demand becomes

$$q_i = [p - p_i + (\alpha_H - E(\alpha)) + t/N]/t.$$

If seller i keeps his price constant—effectively the situation that we capture with our seller-title-price matched dataset—then seller i 's demand increases to $1/N + [\alpha_H - E(\alpha)]/t$. The percentage increase in demand from certification is thus given by

$$[(1/N + [\alpha_H - E(\alpha)]/t) - 1/N]/(1/N) = N[\alpha_H - E(\alpha)]/t.$$

That is, since there is a fixed increase in demand from certification, in proportional terms the impact is increasing in product market competition N .

II. Seller Feedback, Category Quality, and the Value of Certification

We again assume there are two types of sellers: high-types that complete transactions successfully with probability α_H , and low-types that are successful with probability α_L , and that the overall frequency of high-type sellers is ϕ , so the share of low-type sellers is $(1 - \phi)$. Before the seller posts any listings, eBay observes the seller's true quality with probability λ . If eBay observes that the seller is type H , eBay awards it an eTRS badge, so that $B = 1$. In all other cases, $B = 0$. After each transaction, consumers make a public feedback announcement f . Consumers report $f = 1$ if the transaction was good and $f = 0$ if the transaction was bad.

We examine consumer inferences about seller quality for the seller's first two trades. Before deciding whether to buy a product, potential consumers observe the seller's state, (F, T, B) . F is the sum of all prior feedback, T is the number of completed trades, and B is badge status. F , T , and B each take values in $\{0, 1\}$.

We define the badge premium as $\pi(F, T) = E(\alpha|F, T, 1) - E(\alpha|F, T, 0)$. By assumption $E(\alpha|F, T, 1) = \alpha_H$ regardless of F and T . Consumers calculate expected utility $EU = E(\alpha|F, T, B)u - p$, where u is gross value from a perfect transaction and p is price.

We next provide some definitions that simplify our exposition. Let w_L and w_H represent nonnegative constants, and use them to form the probability weights ρ_L and ρ_H . We construct $\rho_a = w_a/(w_H + w_L)$. Clearly $\rho_H + \rho_L = 1$. Holding w_L fixed, ρ_L is decreasing in w_H , while ρ_H is increasing. For the expected value $E(\alpha) = \rho_H\alpha_H + \rho_L\alpha_L$, it follows from the construction of the ρ and w terms that greater values of w_H for fixed w_L imply greater values of $E(\alpha)$.

Our interest is in modeling how consumers' beliefs about seller type evolve over rounds of trade and differ across product categories, which then allows us to compare certification premia across different seller and category traits.

A. Seller Feedback and the Value of Certification

Before the first round of trade.—Before any trade has occurred a seller has $F = T = 0$, and B is equal to 0 or 1. The probabilities of the two possible seller states are $\Pr(0, 0, 1) = \lambda\phi$ and $\Pr(0, 0, 0) = (1 - \lambda)\phi + (1 - \phi)$. The certification rules immediately provide two conditional probabilities: $\Pr(\alpha_H|0, 0, 1) = 1$ and

$\Pr(\alpha_L|0,0,1) = 0$. Applying Bayes' rule, the additional conditional probabilities are:

$$\Pr(\alpha_H|0,0,0) = \frac{(1-\lambda)\phi}{(1-\lambda)\phi + (1-\phi)} \text{ and}$$

$$\Pr(\alpha_L|0,0,0) = \frac{(1-\phi)}{(1-\lambda)\phi + (1-\phi)}.$$

A consumer who sees $B = 1$ immediately infers $E(\alpha|0,0,1) = \alpha_H$. Alternatively, in state $(0,0,0)$ the consumer calculates

$$E(\alpha|0,0,0) = \frac{\alpha_H(1-\lambda)\phi + \alpha_L(1-\phi)}{(1-\lambda)\phi + (1-\phi)}.$$

Note that this expected value can be written as $E(\alpha|0,0,0) = \rho_H \alpha_H + \rho_L \alpha_L$ with appropriately defined probabilities ρ_a . The badge premium $\pi(0,0) = \alpha_H - E(\alpha|0,0,0)$ is positive. This is clear from the probabilities $\rho_H < 1$ and $\rho_L > 0$ in $E(\alpha|0,0,0)$.

Before the second round of trade.—A seller's first round of trade will yield feedback f , so F will be 0 or 1. T automatically increases to 1. The seller's certification state, B , does not change. Possible seller states are $(0,1,0)$, $(1,1,0)$, $(0,1,1)$, and $(1,1,1)$. The various states' probabilities are

$$\Pr(0,1,0) = (1-\alpha_H)(1-\lambda)\phi + (1-\alpha_L)(1-\phi)$$

$$\Pr(1,1,0) = \alpha_H(1-\lambda)\phi + \alpha_L(1-\phi)$$

$$\Pr(0,1,1) = (1-\alpha_H)\lambda\phi$$

$$\Pr(1,1,1) = \alpha_H\lambda\phi.$$

These state probabilities yield the following conditional probabilities:

$$\Pr(\alpha_H|0,1,0) = \frac{(1-\alpha_H)(1-\lambda)\phi}{(1-\alpha_H)(1-\lambda)\phi + (1-\alpha_L)(1-\phi)}$$

$$\Pr(\alpha_L|0,1,0) = \frac{(1-\alpha_L)(1-\phi)}{(1-\alpha_H)(1-\lambda)\phi + (1-\alpha_L)(1-\phi)}$$

$$\Pr(\alpha_H|1,1,0) = \frac{\alpha_H(1-\lambda)\phi}{\alpha_H(1-\lambda)\phi + \alpha_L(1-\phi)}$$

$$\Pr(\alpha_L|1,1,0) = \frac{\alpha_L(1-\phi)}{\alpha_H(1-\lambda)\phi + \alpha_L(1-\phi)}.$$

Note again, for all states with $B = 1$, $\Pr(\alpha_H) = 1$. The relevant expected values are thus

$$E(\alpha|0, 1, 0) = \frac{\alpha_H(1 - \alpha_H)(1 - \lambda)\phi + \alpha_L(1 - \alpha_L)(1 - \phi)}{(1 - \alpha_H)(1 - \lambda)\phi + (1 - \alpha_L)(1 - \phi)}$$

$$E(\alpha|1, 1, 0) = \frac{\alpha_H^2(1 - \lambda)\phi + \alpha_L^2(1 - \phi)}{(1 - \lambda)\alpha_H\phi + \alpha_L(1 - \phi)}.$$

We use these expressions to calculate certification premia, and also to compare the premia to those that obtain prior to the first round of trade. Consider the difference between the no-feedback premium, $\pi(0, 0) = \alpha_H - E(\alpha|0, 0, 0)$, and the premium with one unit of positive feedback, $\pi(1, 1) = \alpha_H - E(\alpha|1, 1, 0)$. This difference is

$$\begin{aligned} \pi(0, 0) - \pi(1, 1) &= E(\alpha|1, 1, 0) - E(\alpha|0, 0, 0) \\ &= \frac{\alpha_H^2(1 - \lambda)\phi + \alpha_L^2(1 - \phi)}{(1 - \lambda)\alpha_H\phi + \alpha_L(1 - \phi)} - \frac{\alpha_H(1 - \lambda)\phi + \alpha_L(1 - \phi)}{(1 - \lambda)\phi + (1 - \phi)}. \end{aligned}$$

We simplify the difference by replacing $w_H = (1 - \lambda)\phi$ and $w_L = (1 - \phi)$, and dividing the numerator and denominator of $E(\alpha|1, 1, 0)$ by α_L :

$$\pi(0, 0) - \pi(1, 1) = \frac{\alpha_H(\alpha_H/\alpha_L)w_H + \alpha_L w_L}{(\alpha_H/\alpha_L)w_H + w_L} - \frac{\alpha_H w_H + \alpha_L w_L}{w_H + w_L}.$$

Next, we set $(\alpha_H/\alpha_L)w_H = \tilde{w}_H$, and note that $w_H < \tilde{w}_H$. Following the properties of w s discussed above, we see that the premium difference is positive, so that the certification premium is greater when no feedback has yet occurred.

While it is possible to also generate comparisons across different F/T ratios, the precise properties would depend on the precise specification of how we model feedback, and in any event our results on the case of $F = T$ serve as the clearest representation of the link between more positive feedback and the value of certification.

B. Cross-Category Comparisons on the Value of Certification

We now augment the model by assuming that some product categories are riskier than others; we further assume that sellers do all of their trade in a single category. We model category risk through ϕ , and say that category j is riskier than k if $\phi_j < \phi_k$. Holding fixed the eBay investigation parameter λ (as is the case in practice), fewer sellers will receive $B = 1$ in a higher risk category because there are fewer high-quality sellers.

We constrain $F = T$ and let $\pi_j(F)$ be the certification premium in a category with high-quality share ϕ_j . We compare premia across categories through the difference

$\pi_j(F) - \pi_k(F)$ for $\phi_j < \phi_k$. If we extrapolate our expression for $E(\alpha_H|1, 1, 0)$ expression above to the general case of F , we may write the relevant difference as:

$$\begin{aligned} \pi_j(F) - \pi_k(F) &= \frac{\alpha_H^{F+1}(1-\lambda)\phi_k + \alpha_L^{F+1}(1-\phi_k)}{\alpha_H^F(1-\lambda)\phi_k + \alpha_L^F(1-\phi_k)} \\ &\quad - \frac{\alpha_H^{F+1}(1-\lambda)\phi_j + \alpha_L^{F+1}(1-\phi_j)}{\alpha_H^F(1-\lambda)\phi_j + \alpha_L^F(1-\phi_j)}. \end{aligned}$$

Replacing $w_{Hk} = \alpha_H^F(1-\lambda)\phi_k$ and $w_{Lk} = \alpha_L^F(1-\phi_k)$, and defining analogous values of w_{Hj} and w_{Lj} , we may write the difference as

$$\pi_j(F) - \pi_k(F) = \frac{\alpha_H w_{Hk} + \alpha_L w_{Lk}}{w_{Hk} + w_{Lk}} - \frac{\alpha_H w_{Hj} + \alpha_L w_{Lj}}{w_{Hj} + w_{Lj}}.$$

Next, multiply the w_{Lk} terms by $\frac{1-\phi_j}{1-\phi_k}$ and the w_{Hk} terms by $\frac{\phi_j}{\phi_k}$. This yields

$$\pi_j(F) - \pi_k(F) = \frac{\alpha_H w_{Hj} \left(\frac{\phi_k}{\phi_j}\right) + \alpha_L w_{Lj} \left(\frac{1-\phi_k}{1-\phi_j}\right)}{w_{Hj} \left(\frac{\phi_k}{\phi_j}\right) + w_{Lj} \left(\frac{1-\phi_k}{1-\phi_j}\right)} - \frac{\alpha_H w_{Hj} + \alpha_L w_{Lj}}{w_{Hj} + w_{Lj}}.$$

Notice that we have eliminated w_{Hk} and w_{Lk} . We now multiply the top and bottom of the first term by $\left(\frac{1-\phi_j}{1-\phi_k}\right)$ to get

$$\pi_j(F) - \pi_k(F) = \frac{\alpha_H \tilde{w}_{Hj} + \alpha_L w_{Lj}}{\tilde{w}_{Hj} + w_{Lj}} - \frac{\alpha_H w_{Hj} + \alpha_L w_{Lj}}{w_{Hj} + w_{Lj}}$$

$$\text{with } \tilde{w}_{Hj} = w_{Hj} \left(\frac{\phi_k}{\phi_j}\right) \left(\frac{1-\phi_j}{1-\phi_k}\right).$$

The final remaining step is to show $\tilde{w}_{Hj} > w_{Hj}$, which follows directly from the assumption that $\phi_j < \phi_k$, since both $\left(\frac{\phi_k}{\phi_j}\right) > 1$ and $\left(\frac{1-\phi_j}{1-\phi_k}\right) > 1$. Thus, certification's value will be greater in markets where certification is rarer due to category risk.

REFERENCES

- Arrow, Kenneth J.** 1972. "Gifts and Exchanges." *Philosophy and Public Affairs* 1 (4): 343–62.
- Bajari, Patrick, and Ali Hortaçsu.** 2004. "Economic Insights from Internet Auctions." *Journal of Economic Literature* 42 (2): 457–86.
- Becker, Bo, and Todd Milbourn.** 2011. "How Did Increased Competition Affect Credit Ratings?" *Journal of Financial Economics* 101 (3): 493–514.
- Bolton, Gary, Ben Greiner, and Axel Ockenfels.** 2013. "Engineering Trust: Reciprocity in the Production of Reputation Information." *Management Science* 59 (2): 265–85.
- Brown, Jennifer, and John Morgan.** 2006. "Reputation in Online Markets: Some Negative Feedback." <http://faculty.haas.berkeley.edu/rjmorgan/reputation%20in%20online%20markets.pdf>.
- Dini, Federico, and Giancarlo Spagnolo.** 2009. "Buying reputation on eBay: Do recent changes help?" *International Journal of Electronic Business* 7 (6): 581–98.
- Dranove, David, Daniel Kessler, Mark McClellan, and Mark Satterthwaite.** 2003. "Is More Information Better? The Effects of 'Report Cards' on Health Care Providers." *Journal of Political Economy* 111 (3): 555–88.
- Dranove, David, and Ginger Zhe Jin.** 2010. "Quality Disclosure and Certification: Theory and Practice." *Journal of Economic Literature* 48 (4): 935–63.
- Einav, Liran, Chiara Farronato, Jonathan D. Levin, and Neel Sundareshan.** 2013. "Sales Mechanisms in Online Markets: What Happened to Internet Auctions?" National Bureau of Economic Research (NBER) Working Paper 19021.
- Einav, Liran, Theresa Kuchler, Jonathan Levin, and Neel Sundareshan.** 2015. "Assessing Sale Strategies in Online Markets using Matched Listings." *American Economic Journal: Microeconomics* 7 (2): 215–47.
- Elfenbein, Daniel W., Ray Fisman, and Brian McManus.** 2012. "Charity as a Substitute for Reputation: Evidence from an Online Marketplace." *Review of Economic Studies* 79 (4): 1441–68.
- Elfenbein, Daniel W., Raymond Fisman, and Brian McManus.** 2015. "Market Structure, Reputation, and the Value of Quality Certification: Dataset." *American Economic Journal: Microeconomics*. <http://dx.doi.org/10.1257/mic.20130182>.
- Forbes, Silke J., Mara Lederman, and Trevor Tombe.** 2015. "Quality Disclosure Programs and Internal Organizational Practices: Evidence from Airline Flight Delays." *American Economic Journal: Microeconomics* 7 (2): 1–26.
- Ho, Daniel E.** 2012. "Fudging the Nudge: Information Disclosure and Restaurant Grading." *Yale Law Journal* 122 (3): 574–688.
- Hui, Xiang, Maryam Saeedi, Zeqian Shen, Neel Sundareshan.** 2013. "From Lemon Markets to Managed Markets: The Evolution of eBay's Reputation System." https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=IIOC2014&paper_id=465.
- Jin, Ginger Zhe, and Phillip Leslie.** 2003. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." *Quarterly Journal of Economics* 118 (2): 409–51.
- Jin, Ginger Zhe, and Phillip Leslie.** 2009. "Reputational Incentives for Restaurant Hygiene." *American Economic Journal: Microeconomics* 1 (1): 237–67.
- Lewis, Gregory.** 2011. "Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors." *American Economic Review* 101 (4): 1535–46.
- Luca, Michael.** 2011. "Reviews, Reputation, and Revenue: The Case of Yelp.com." Harvard Business School Working Paper 12-016.
- Ramanarayanan, Subramaniam, and Jason Snyder.** 2012. "Reputations and Firm Performance: Evidence from the Dialysis Industry." <http://www.anderson.ucla.edu/faculty/subramaniam.ramanarayanan/Papers/Dialysis%20submission.pdf>.
- Roberts, James W.** 2011. "Can Warranties Substitute for Reputations?" *American Economic Journal: Microeconomics* 3 (3): 69–85.
- Saeedi, Maryam.** 2012. "Reputation and Adverse Selection: Theory and Evidence from eBay." <http://www.econ.umn.edu/saeedi/saeediJMP.pdf>.
- Salop, Steven C.** 1979. "Monopolistic Competition with Outside Goods." *Bell Journal of Economics* 10 (1): 141–56.
- Tadelis, Steven, and Florian Zettelmeyer.** 2011. "Information Disclosure as a Matching Mechanism: Theory and Evidence from a Field Experiment." *American Economic Review* 105 (2): 886–905.
- Wimmer, Bradley S., and Brian Chezum.** 2003. "An Empirical Examination of Quality Certification in a 'Lemons Market.'" *Economic Inquiry* 41 (2): 279–91.
- Xiao, Mo.** 2010. "Is Quality Certification Effective? Evidence from the Childcare Market." *International Journal of Industrial Organization* 28 (6): 708–21.