Lifting Growth Barriers for New Firms
Evidence from an Entrepreneurship Training Experiment with Two Million Online Businesses*

Yizhou Jin† Zhengyun Sun‡
(Job Market Paper)

October 31, 2020

Latest version here

Abstract

Expansion of e-commerce presents new opportunities for small and medium enterprises (SMEs) to enter broader markets at lower costs, but the new entrants face barriers to growth after entry. To facilitate the new entrants to overcome these barriers, we implement a training program as a randomized controlled experiment with over two million new sellers on a large e-commerce platform. The training focuses on practical skills specific to online business operations. Treated new sellers with access to the training earn higher revenues. These sellers improve marketing skills and attract more consumers to their online stores. Leveraging detailed consumer-seller matched search and browsing data, we find that consumers have higher purchase probability overall when they encounter new sellers regardless of treatment status. In the cases of purchases, consumers choose treated new sellers over incumbents; moreover, doing so does not lower the quality of their purchases. We use a structural model to characterize consumer demand and recover sellers' underlying quality. Both treated and control new sellers have higher quality compared to incumbents. The training increases new sellers' likelihood of being encountered by consumers, which improves the matching quality between consumers and sellers. The counterfactual exercise shows that training leads to higher consumer surplus and the platform's total sales due to market expansion. The platform could increase profits in both the short and the long run because of the training.

Keywords: e-commerce, platforms, business training, growth barriers
JEL codes: L26, L81, M53, O12, O17

*Sincere thanks to our advisors Jie Bai, Emily Breza and Michael Kremer for their extensive advice and guidance. We thank the product development team at the unnamed e-commerce platform for making the experiment possible. For especially valuable comments and discussion, we thank Long Chen, Shawn Cole, Asim Khwaja, Gabriel Kreindler, Song Ma, Ariel Pakes, Gautam Rao, Yulu Tang, Yuan Tian, Christopher Woodruff and various seminar participants at the Development Lunch and Seminar at Harvard. We also thank our data contacts and providers for making this project possible. They include the unnamed e-commerce platform, Carolina Harvey. Ability to publish is not contingent on results. Please email carolina_harvey@harvard.edu for questions about the data usage agreements.

†Haas School of Business, UC Berkeley. Email: jyz@berkeley.edu

‡Department of Economics, Harvard University. Email: zsun@g.harvard.edu

1 Introduction

The growth of e-commerce is accelerating in many developing countries. In China, e-commerce sales grew at an average annual rate of 25 percent for the past five years (Ministry of Commerce, 2020). In 2019 alone, e-commerce sales grew by 32 percent in India and 25 percent in Mexico (Lipsman, 2019). The expansion of e-commerce provides particularly exciting new opportunities for small and medium enterprises (SMEs) because of reduced entry costs and extended market access (World-Bank, 2016; Lehdonvirta et al., 2019). However, many challenges still remain for new entrants to survive and grow after entry. In particular, these new entrants need to learn about online business operations, which requires skills such as internet marketing and customer management that are different from running offline businesses. Moreover, as Bai et al. (2020) have recently shown, sellers on a cross-border e-commerce platform need to overcome sizable search and information frictions in order to grow.

Lifting growth barriers for promising new entrants could be beneficial for the e-commerce platforms. As new sellers bring more varieties to the market and increase competitive pressure on incumbents, consumers stand to have a better experience overall. The platform has incentives to be more proactive and to support the promising new sellers, since these actions might yield long and short-term benefits. In the long-run, a better market environment allows the platforms to attract and retain more consumers and sellers, which coincides with the platform's profit-maximizing goal. In the short-run, the platform could also benefit directly if sellers earn higher revenues and invest more in marketing on the platform. To support the new entrants, one approach that the platform could adopt is to ensure that new sellers master the basic skills of online business operations so that a knowledge gap does not hinder their growth. In this paper, we study the impacts of one of such efforts by a prominent e-commerce platform: a large-scale business training program designed to help new sellers overcome growth barriers.

The e-commerce platform's efforts to promote the growth of small businesses with training follow the footsteps of many predecessors. However, despite previous efforts, the effects of the training on relevant market participants are still ambiguous. For supported firms, typical business training that teaches best business practices has mixed impacts on profits and growth (McKenzie and Woodruff, 2014). While Management consulting is effective, its high costs make it difficult to scale up (Bloom et al., 2013; Bruhn et al., 2018). For non-supported firms, scarce evidence shows that spillover could be limited. McKenzie and Puerto (forthcoming) varied the treatment intensity of a training intervention at market level and found no significant spillover on competitors. Apart from business training, some empirical studies evaluate the spillover effects of firm subsidies (Rotemberg, 2019), credit access (Banerjee and Duflo, 2014) and microfinance (Banerjee et al., 2015). However, what is relatively understudied is the impact of interventions to support small businesses on consumers in the markets.

In this paper, we implemented a business training intervention as a randomized controlled experiment with over two million new sellers on a large e-commerce platform to answer the following questions. First, can the training lift growth barriers for new sellers on the platform? If so, through what channels? Second, how does the training affect consumers' experiences on the platform? Third, what are the welfare implications of the training on new sellers, incumbents and consumers?

The e-commerce platform with which we collaborate hosts millions of consumers and sellers. Sellers on the platform are mostly retailers that offer diverse products. We implement the training program at scale, taking advantage of the close to zero marginal dissemination costs online. In contrast to typical business training that teaches generic best business practices, our training program focuses on practical online business operation and marketing skills. We randomly assign access to the training program when new sellers register on the platform. To date, over two million sellers received access. In our study cohorts, 24.9 percent of all the registered new sellers have access and 24.1 percent of sellers with access took up the training in the following nine-month period.

To study the impacts of the training on new sellers, we leverage random assignment of the training access and compare the performance of treated and control new sellers. Rich administrative data also allows us to investigate the impacts on sellers' product offering, marketing and customer service. Next, to evaluate the impacts of the training on consumers, we use rich consumer-seller matched browsing data to recover set of sellers that consumers visited when they search for specific products on the platform and exploit variations in the search results. Lastly, we use a structural model to characterize consumers' demand and recover the rules that the platform uses to match consumers and sellers in the search results. With the model, we decompose the welfare impacts of the training on new sellers, incumbents and consumers.

We find that the training changes the experience of new sellers, incumbents and consumers on the platform. First, training lifts new sellers' growth barriers as treated new sellers earn higher revenues. Compared to new sellers in the control group, new sellers with access to the training earn 1.7 percent higher revenues. Using random assignment of the training as the instrument, we find that sellers who participate in the training earn 6.6 percent higher revenues. The revenue gains occur mostly because treated sellers attract more consumers to their sites as their marketing skills improve. Specifically, these sellers participate more in pay-per-clicks ads and promotional events to attract consumers. In addition, treated new sellers improve their customer service quality as they adopt more supplementary services such as the AI assistant to handle customer inquiries. However, we do not find that treated new sellers have significantly higher average purchase probability among visitors or more positive customers' ratings than new sellers in the control group.

As the training helps new sellers to accumulate customers, consumers are more likely to encounter these sellers on the platform. Overall, consumers have higher purchase probability when they visit new sellers - treated or control - than when they visit incumbents. This result holds after controlling for consumer, search keyword and search effort specific effects. To confirm it is the new sellers who are driving the results, we check whom the consumers choose if they do make purchases and find that they indeed choose treated new sellers over incumbents. In the meantime, we do not find adverse effects on the quality of purchases: consumers are no more likely to request returns or refunds when they purchase from new sellers, while they are as likely to make repeat purchases. Therefore, the training enables promising, higher quality new sellers to interact more often with consumers, which benefits the consumers because the matching quality is higher with these new sellers than with incumbents.

Based on the reduced-form results, we build a structural model to characterize consumer demand and recover the rule that the platform uses to match sellers and consumers. With the model, we use variations in consumers' choice probabilities to recover underlying sellers' quality. Among the set of new sellers and incumbents that consumers visited, both treated and control new sellers have significantly higher underlying quality than incumbents. The difference in consumers' purchase probability when they encounter different types of sellers suggests that a main friction in the market is that high quality new sellers are not being encountered by consumers often enough. We conduct a counterfactual exercise to evaluate the welfare impacts of the training. We remove training participants' access to the training by lowering the participants' chances to be found by consumers in the matching[1]. Doing so causes a 0.1 percent decline in consumer surplus and in total sellers' revenues, since consumers are less likely to interact with higher quality new sellers. The revenue drop is driven by consumers making fewer purchases as they interact with fewer new entrants in the absence of the training, while the impacts on the market reallocation between incumbents and new sellers is limited.

Our study relates to several strands of literature. First, empirical works studying growth barriers and firm dynamics for new entrants have recently shifted their focus to demand side frictions. For offline firms, previous research highlights barriers to growth due to the lack of initial market access (Atkin et al., 2017), slow customer accumulation (Foster et al., 2016; Piveteau, 2016) and the uncertainty in learning (Arkolakis et al., 2018; Berman et al., 2019). Overall, our study most closely relates to the work done by Bai et al. (2020). The authors highlight that the demand-side search frictions limit high quality sellers' growth in

[1] Share of training participants removed is consistent with what we estimated in the reduced form results.

a cross-border e-commerce market. Our paper confirms that such frictions are limiting for new entrants' growth in e-commerce and experimentally verifies that the business training that closes sellers' knowledge gap and improves their marketing skills could be an effective strategy not only to lift growth barriers but also to improve consumers' experience on the platform[2].

Second, we contribute to an extensive literature on business training intended to help SMEs in the developing world. McKenzie and Woodruff (2014) reviews this literature and finds mixed results on the effectiveness of training for offline firms. Our experiment shows that training is a low-cost way to lift growth barriers for new entrants in online markets. The design of the online training borrows from previous success stories in the literature, incorporating large-scale customization (Bloom et al., 2013; Bruhn et al., 2018) and rule of thumb style tutorials (Drexler et al., 2014). On specific mechanisms, our finding that better marketing could facilitate the growth of new entrants echoes the findings in Anderson et al. (2018), where the authors show that a business training that teaches marketing skills paves a growth-focused pathway to profits. More broadly, marketing and building customer capital is an important mechanism for growth in many settings (Gourio and Rudanko, 2014; Fitzgerald et al., 2016). For broader implications of efforts to support SMEs, some recent empirical studies examine "experimentation at scale" (Muralidharan and Niehaus, 2017) to evaluate effects on non-treated market participants. For business training, Calderon et al. (2020) randomize access to training at village level in Mexico and find no spillover effects partially due to small sample size. Recent work by McKenzie and Puerto (forthcoming) uses a two-stage experimental design where the authors randomize the intensity of a business training intervention at market level and then randomize individual businesses' access to the training within each market in Kenya. Three years after the training, not only did treated firms earn higher profits, but their success did not come at the costs of their competitors, as the market expanded in terms of sales. Instead of varying treatment intensity across markets, we contribute to this line of work by pointing out a novel channel of welfare gains with unique consumer-seller matched data. In an online market with friction, business training could promote high quality sellers in the matching between consumers and sellers, thereby improving both consumers' experience and market efficiency.

Third, we speak to the recent literature that examines the expansion of the digital economy (Goldfarb and Tucker, 2019) and the line of work on the roles of platforms as regulators of the various markets they host. As technology such as high-speed internet creates

[2] A line of literature investigates consumers' search frictions in various online markets and evaluates the platform's design to improve search efficiency. Some examples of empirical works include Fradkin (2015) (Airbnb), Dinerstein et al. (2018) (eBay), Horton (2014) (labor market), Ursu (2018) (Expedia) and Chen and Wu (2020) (AliExpress).

business opportunities (Hjort and Poulsen, 2018), many challenges remain for newcomers. For example, Couture et al. (2018) show the vastly heterogeneous consumption side responses and lack of supply side reactions as e-commerce penetrates rural retail markets. While the concerns over e-commerce platforms' market power are looming, many recent studies show how e-commerce could foster competition, improve efficiency and boost consumer welfare[3]. Outside e-commerce, peer-to-peer platforms lower transaction costs and reduce search frictions[4]. We add to this discussion by emphasizing the importance of platforms' interventions on reducing frictions and maintaining a more equitable, competitive environment for market participants (Tadelis, 2016; Hui et al., 2016; Cui et al., 2020). Interventions such as the business training helps the platform to achieve its profit-maximizing goal and improve sellers and consumers' experience.

The rest of the paper is organized as follows. In section 2, we first describe features of the e-commerce platform and then introduce the training intervention and the experimental design. In section 3, we examine the impacts of the training on new sellers. Then, in section 4, we analyze the impacts of the training on consumers. With the reduced-form results, we build a structural model to decompose the welfare implications of the training in section 5. Lastly, in section 6, we conclude.

2 Business Training on the E-commerce Platform

2.1 Sellers on the E-commerce Platform

In this paper, we partner with a leading e-commerce platform in China. The platform hosts millions of active sellers and consumers and total sales on the platform represents a sizable share of all domestic retail sales. Sellers on the platform offer diverse products. Some of the most popular sectors include clothing, cosmetics, home appliances, consumer electronics and food. Vast majority of sellers on the platform are retailers who source their products from manufacturers or wholesalers. Unlike Amazon, third-party sellers generate dominant share of sales on the platform. The platform earns revenues from these sellers by offering advertisements, charging commissions and selling supplementary services. Platform's reliance on the third-party sellers implies that ensuring these sellers can thrive on the platform is in alignment with the platform's profit-maximizing goal. Therefore, the platform is motivated to implement policies and programs that bring in more third-party sellers and foster their growth after entry. Couture et al. (2018) investigates one of such

[3] Some empirical analysis of welfare impacts of e-commerce include Brynjolfsson et al. (2003), Einav et al. (2017) and Jo et al. (2019).

[4] Empirical studies situate in different platform markets and show that while frictions still exist, platforms still have the power to improve efficiency using various algorithms and mechanism, see for example Cohen et al. (2016), Farronato and Fradkin (2018) and Ellison and Ellison (2018).

efforts where the platform brings e-commerce to rural villages.

Comparing to starting an offline business, becoming a registered seller on the platform requires considerably lower monetary and effort costs. To register as a sole proprietor, a potential new seller only needs to complete the authentication with a national ID or a formal business registry but does not need to pay any registration fees. Except for several regulated sectors, the platform does not ask for certificates and charges zero commission fees [5]. These sole proprietors make up for roughly 97% of all active sellers on the platform. Most of the active sellers are highly dedicated. Running the e-commerce businesses is a full-time job and the main source of income for these sellers. According to an online survey with selected new sellers in the training sample, majority of the respondents state that their intend to operate the online store as their main jobs. Appendix C discusses more details of the survey.

Despite the easy registration, sellers face growth barriers after entry. First of all, posting and selling products on the platform come at additional monetary and effort costs[6]. In some cases, an inexperienced new seller could spend more than 30 minutes to post a product in order to provide accurate descriptions and pictures that meet platform's requirements. After posting products, attracting visitors to the sites is the prerequisite to grow for both new sellers and incumbents. The platform uses sophisticated algorithm in search and recommendation to match consumers with right sellers in order to achieve the most efficient outcomes. Sellers can influence the results by actively engaging in complex strategies to compete for consumers[7]. The most common strategies are purchasing pay-per-click ads, participating in promotions that the platform regularly organizes and recruiting celebrities to do marketing on social media. In the most cases consumers search for products rather the sellers and sellers compete for better rankings in each search session. Such competition in a search session is close to zero-sum[8]. Advertising and promotion will directly influence search rankings, but social media based marketing operates through a different channel.

[5]The regulated sectors include food, drug, medical equipment, cigarettes, liquor, infant formulas, and other products that are subject to public health and safety concerns. Our analysis focus on the C2C ("consumer-to-consumer") sellers. The platform also hosts a small number of "business-to-consumer" (B2C) sellers. These sellers are formally registered, have brands and completed formal applications to operate on the platform. Consumers can access these two types of sellers' sites on the most popular app the platform offers, but B2C sellers have special demarcation for their status and get preferential treatment in search rankings. B2C sellers are also much larger and some of these sellers are internationally recognized brands.

[6]Before posting products on the platform, sellers need to put down small deposits as "consumer protection fees" for potential dispute resolutions. Exact requirements differ by sectors, typically ranging from 0 to 5000 RMB. Sellers can still list their products on the platform but their products will get much lower rankings in search results and will not be promoted in other channels

[7]Interviews with multiple sellers on the platform suggest marketing spending could account for significant share of the operating costs. Larger sellers invest even more heavily than small sellers.

[8]Currently over 90% of consumers accessing the platform are from mobile device rather than from the web. Therefore, competition for ranking is more intense because of limited space per screen on the mobile device. On the other side, it is hard to define an obvious page break in the search results.

The sheer number of competitors on the platform and the intense competition between these sellers for consumers' attention indicates that marketing is a crucial component of online business operation.

2.2 Business Training

We collaborate with the platform to implement an online business training program as a randomized controlled experiment. The low dissemination costs for online intervention make it feasible for the training to reach a large number of sellers. The program was officially launched on May 6, 2019 and is available since then. Over two million new sellers have access to the training by June 2020. The training is a standalone program independent from other operations of the platform. In particular, participation and performance in the training does not affect how the platform matches consumers with sellers in the searches.

The platform partners with professional e-commerce service providers to design the training. In contrast to typical business training that teaches generic best business practices, this training focuses on specific challenges of running e-commerce businesses in order to help new sellers better navigate the platform market. The training materials are organized as sequences of tasks and each task tackles a specific challenge. In the training, the platform uses administrative data to dynamically match sellers with the most appropriate tasks based on sellers' performance and actions. Each task uses a combination of tutorials, Q&A forums and webinars to deliver recommendations.

Participating in the Training New sellers can access the training on the Official Seller's Portal app where the training module appears as a widget on the front page (see figure 1 panel A)[9]. The official app is essential for sellers to manage their stores and to communicate with the platform. Therefore, dedicated sellers need to install and use the app regardless, so that they do not need to invest additional efforts to access the training.

Each task singles out an area of improvement and sellers can choose which of the tasks to try (see figure 1 panel B). The training tasks are associated with specific performance metrics along with corresponding triggering conditions. For example, a new seller triggers the task "attract more visitors to your store" if the number of visitors she had over the past 30 days is below 40th percentile among sellers in her sector. An algorithm examines sellers' performance and assigns the most relevant tasks based on the performance and the triggering conditions of the tasks.

After taking up the tasks, sellers can access detailed tutorials written by professional e-commerce service providers that the platform collaborates with. Each task has an associated Q&A forum where sellers can directly reach out to the authors of the tutorials. Some service providers also offer live-streamed webinars to directly communicate with

[9]The app is available on all major operating systems and has a web version.

the sellers. Tasks have varying completion time based on their difficulties, usually ranging from three days to a month. Reaching the pre-specified targets marks the completion of the tasks. Sellers need to take specific actions or out-perform other sellers in their sectors. For example, a seller need to have number of visitors above 60th percentile among sellers in her sector during the past 30 day to complete the "attracting more visitors to your store" task. Sellers earn short-term free accesses to certain supplementary services for each task they complete. These supplementary services support routine online business operations[10]. The monetary value of the short-term access to the services range from $5 to $10. Sellers are unaware of the rewards before taking up the first tasks. If a seller fails to complete a task, she can always make additional attempts later[11].

Content of the Training The tasks in the training cover three major areas of online business operations: basic setup, marketing and customer service[12]. The first type of tasks focuses on teaching new sellers how to set up online stores without running into pitfalls or violating the platform's rules. A typical task in this category teaches new sellers how to post products on the platform. The tutorial of the task contains step-by-step guide to ensure that sellers follow the platform's rules and to help the products get better search rankings[13]. Taking up these tasks might increase sellers' likelihood to set up their online stores and post products to sell. The second type of tasks addresses challenges to attract consumers by teaching relevant marketing techniques. Typical tasks in this category teach sellers ways to improve quality of their product titles in order to get better search rankings, methods to select more suitable keywords used in pay-per-click ads and techniques to take advantage of hundreds of sales events the platform regularly organizes. Sellers taking up these tasks could be more active in marketing and could improve marketing skills. The last type of tasks focuses on improving sellers' customer service quality. In this category, typical tasks introduce sellers to many supplementary tools that the platform offers to help sellers better manage their stores. In one task, the tutorial teaches sellers how to set up an AI assistant to answer consumers' inquiries in timely fashion. Adoption of the tools could help improve customer satisfaction and service quality.

The goal of the training is to help new sellers to better navigate the complex online business environment. The training puts heavy emphasis on pushing sellers to stay active in

---

[10] As an example, sellers can access a program that allows sellers to print many customized shipping labels with one-click.

[11] The training module does not explicitly state that there is no consequence of not completing the tasks, which might deter some sellers worrying about potential negative consequences. We do not have empirical supports for the direction of selection.

[12] Appendix table A12 provides a list of tasks with detailed contents and classifications.

[13] Complexity of the product management system on the platform makes posting product a non-trivial task. Each posting requires sellers to describe characteristics of the products in great details. Such information is important inputs for the search algorithm. An example of a trick is at least one of product pictures should have white background in order to get promotion in non-search channels.

the market and on attracting more visitors with better marketing. The focus on customer acquisition echoes earlier findings that demand-side frictions could be a main growth barrier new sellers face in the e-commerce market. While some techniques are relevant only for this platform, many marketing and customer management skills can be easily transplanted when operating other online businesses. Training does not cover more generic business practices often cases mentioned in the literature such as managing supply chain, finance and personnel (Bloom and Van Reenen, 2007).

For some sellers, the training program helps close their knowledge gap by teaching previously unknown techniques. For the rest, the information is already available from other sources. The training then functions as a well-structured reminder to remind new sellers about what should be done at certain stages of their growth trajectory.

2.3 Experimental Design and Implementation

We design and implement the training intervention as a randomized controlled experiment where we randomly offer newly registered sellers access to the training. The access is assigned immediately after the sellers complete the registration. The timing of the assignment limits the available baseline information to variables collected during the registration and actions taken on the first day of entry[14]. Moreover, because the platform dynamically matches tasks and sellers based on sellers' performance, we are unable to randomize the assignment at the task level.

The experiment officially started on May 6, 2019 and has been on-going since[15]. In 2019, about 25,000 new sellers registered everyday. From May 6 to October 28, 2019, during the first phase of formal roll-out of the training intervention, we randomly selected 25% of new sellers to access the training. Later on, we expanded the share of new sellers to access the training to 35% on October 29 and to 90% on December 26. Figure 2 summarizes the timeline of treatment assignment. By June 2002, over two million new sellers have access to the training program. For the empirical analysis, we focus on the cohorts of sellers registered between May 6 and August 15, 2019 in order to analyze long-term impacts.

2.4 Training Take-up

Among sellers in the treatment group, 44.6% of the sellers browsed the training during the first month of entry. In the end, 24.1% of treated sellers took up at least one task and 12.6% completed at least one task in first nine months after registration. None of the sellers in the control group took up any tasks. The adoption rate is typical for online products,

[14]Information collected during the registration process include sellers' type (registered as individual or business) and locations. For individual sellers, we also know their gender.

[15]The training program went into testing in April 2019, during which about 1% of new sellers received the access.

but is considerably lower comparing to other training programs offline (McKenzie and Woodruff, 2014)[16].

Sellers are most active to participate in the training during the first month of entry, partially because the content of the training is most relevant for the brand new entrants. Altogether 49.9% of tasks were taken-up during the first month of entry. During this time, sellers are more likely to pick tasks basic store setup and customer acquisition than those on customer management (figure A1). Among sellers who eventually posted products, 7.0% of sellers in the treatment group took up tasks prior to posting products. Tasks on basic setup related challenges have higher completion rate at 58.0%, comparing to the average rate at 37.5%.

Although sellers have access to the training for at least six months after the registration, retention rate declined relatively fast over time. 13.9% of treated sellers took up tasks during the first month, only 4.0% of sellers continued to do so in the third month[17]. Conditioning on taking up some tasks in the previous month, around 23% of sellers will take up more tasks in the following month. Figure 3 shows share of sellers who browsed, took up and completed the tasks unconditionally (left panel) or conditioning on having done so in the previous month (right panel).

Columns 4 to 6 of table 1 summarize characteristics of treated sellers by whether or not they participate in the training. Comparing to sellers in the treatment group who did not participate, sellers who took up some tasks are slightly more likely to be registered as firms but are less likely to be female[18]. Sellers from more economically developed southern coastal provinces are much more likely to participate and they took up more tasks conditioning on participation[19]. In fact, sellers from the southern provinces play the dominant roles on the platform. We hypothesize that sellers who post products on the very first day of entry to be better prepared. The early-movers are more likely to become training participants but conditioning on participation, they do not take up more tasks.

3 Effect of the Training on Sellers

In this section, we discuss the impacts of the training on alleviating new sellers' growth barriers. We analyze the overall impacts of accessing the training on sellers' performance and their strategies.

[16]Take-up rates for typical offline training programs are not perfect, usually in the range of 50 to 90%. Comparing to the offline setting, costs of taking up the training is much lower, but perceived benefits might also be low especially since many competing training programs exist in the market.

[17]Many sellers exit the platform after a month. 89.7% of sellers have visitors during the first month but only 42.2% do during the third month. The rates are similar for sellers in the treatment and control group.

[18]As mentioned, sellers can either register with a national ID card (as individual) or with a formal business registry (as firm).

[19]The coastal provinces are Guangdong, Jiangsu, Zhejiang, Shanghai and Fujian.

3.1 Data: the New Seller Panel

Our main data source is the administrative data that the platform collects on sellers' performance, strategies, characteristics and their participation in the training[20]. As mentioned, we focus on the cohorts of new sellers registered between May 6 and August 15, 2019. On average, 22,230 new sellers registered on the platform during the sample period each day.

We require newly registered sellers to log in to the Official Seller's Portal app at least once within the first seven days of registration and have completed the entire registration process to be included in the final sample[21]. Since sellers will not find out whether or not they have access to the training program prior to the first login to the Seller's Portal, the login requirement does not induce selection. However, the training could potentially affect frequency of subsequent logins[22]. The full sample consists of 712,118 sellers out of which 177,026 (24.8%) sellers are randomly assigned to access the training. We check the balance of the treatment assignment using sellers' characteristics collected during the registration, i.e. their types and locations (column 1 to 3 in table 1). All the characteristics are balanced across treatment and control except for sellers' registration type, where treated sellers are more likely to be firms. Among sellers registered as individuals, 45.3% are females. Gender distribution is balanced, so do the locations at both province and city level.

We construct a balanced monthly panel with all sellers in the final sample. The panel spans the subsequent nine months following the registration. We define each month as a 30-days period relative to the date of entry. The entry day is also the day of treatment assignment. We collect sellers' performance measure on revenues, number of visitors and conversion rates. A seller's conversion rate is the share of visitors who make purchases. Conversion rate is the most commonly used metric for measuring efficiency. For example, the platform uses the conversion rate to evaluate the efficiency of its search and recommendation algorithm. We also collect measures of sellers' quality including their customer ratings (on accuracy of product description, customer service and logistic), likelihood for consumers to request refunds or returns and frequency of violating platform's rules. On sellers' strategies, we observe their product offerings, pricing level, marketing and customer service. We do not observe actual spending on marketing. Instead, we observe some proxies for sellers' engagement, such as number of products participating in pay-per-click

[20]Similar to other studies (Zhang et al., 2019) on online businesses, our data only includes activities that are observable online. Specifically we do not have costs or other offline, supply side information. Therefore our measure of performance would be revenues rather than profits. Moreover, we do not have access to sellers' outcomes on other platforms if they are multi-homing. Multi-homing is more common for larger, more established sellers than small sellers. In our sample this possibility might not be a significant concern for small new sellers. Survey evidence previously mentioned suggest that vast majority of the sellers have very limited offline presence and operate solely online.

[21]Sellers would have their identification related information and locations recorded if they completed the registration.

[22]We empirically test the impacts of the treatment on login but do not find any significant differences.

ads. For sellers with treatment access, we obtain their entire history of interactions with the training program on tasks take-up and completion[23].

In the final sample, 35% of registered sellers have never posted any products to sell. For sellers who posted products or earned revenues, we obtain their affiliated sectors[24]. For sellers without product postings and sellers who exit the market, we replace missing values with zero for outcome measures such as the number of visitors and revenues[25]. Conversion rates are undefined if the sellers attract no visitors. Similarly, the quality measures such as customer ratings are undefined if the sellers do not have any orders. We left these variables as missing. Distribution of the outcome variables such as number of visitors and revenues are extremely skewed. For the main analysis we convert these variables to log scale[26].

3.2 Impacts of the Training

Overall Impacts We first evaluate the overall intent-to-treat (ITT) effect with the new seller panel. For seller i during (relative) month m belonging to an entry-date cohort c with affiliated sector s, we run the following specification:

$$Y_{imcs} = \beta Treatment_i + \alpha_m + \alpha_c + \alpha_s + \epsilon_{imcs} \tag{1}$$

$Treatment_i$ is an indicator for having access to the training, α_m, α_c and α_s are month, cohort and sector fixed effects. Standard errors are clustered by seller.

Table 2 presents the estimated results on sellers' performance. Access to the training leads to a 5% increase in the likelihood of earning revenues as well as 1.7% increase in revenues earned[27]. Since revenues are zero for 60.5% of sellers-month pairs, the unconditional average revenue is close to zero. Restricting the sample to seller-month pairs with positive revenues (column 3 of table 2), sellers in the treatment groups earn 2.6% higher

[23] As described in section 2.2, the assignment of tasks is individualized with temporal variations. We do not keep track of the tasks that were assigned to sellers on daily basis.

[24] Sellers' sectors are determined by products they posted and sold. Therefore, sellers will not have a sector affiliation if they do not post anything. Moreover, a significant share of sellers are labeled as selling second-hand products, which are treated differently from in the search results. We group sellers without sectors, sellers selling second hand products and sellers selling unclassified products together. Since sellers could change their sectors, we use the first sectors that the sellers identify with as their affiliated sectors.

[25] Platform automatically remove a registered seller from the platform if the seller does not have any active product posting over the past four weeks. Notice that there are no requirements on number of visitors attracted or revenues made. When removed, the seller's site is inaccessible and the platform stops collecting data. Sellers have the option to re-open their store, at which point the platform will start to collect the information again under the same ID.

[26] 59.5% of seller-month observations have no visitors. To avoid dropping most of the sample, we add one to the number of visitors, revenues and other performance measure before taking logs.

[27] Table A1 presents the treatment effect on revenues in levels with different winsorization thresholds. Besides the specification with raw revenues, all the estimated effects are positive, but not all significant. The results are very sensitive to extreme values at the top of the distribution. Because of the content of the training, we do not expect the training could have meaningful impacts on these top sellers.

revenues. Treated sellers earn higher revenues because they attract more visitors to their sites. Treated sellers attracted 1.3% more visitors to their sites (column 4) and they have 0.8% more consumers making purchases (column 6). Conditioning on having some visitors, treated sellers attract 2.4% more visitors (column 5) and consequently have 1.8% more consumers making purchases (column 7). However, we do not find a significant improvement on treated sellers' conversion rates as shown in column 8. We use consumer side data to explore the conversion in section 4. Figure A2 presents the quantile treatment effect on log revenues separately for each month. The variations of treatment effects by different quantiles are small. The treatment effect is slightly larger for sellers in the middle of the revenue distribution, especially for sellers who are on the edges of earning revenues.

Table A5 summarizes the impacts of training on sellers' observed quality metrics. We do not find treated sellers to have significantly higher customer ratings comparing to sellers in the control group. For all three types of the ratings, namely accuracy of product descriptions, customer service and logistics, treated sellers obtain slightly higher scores than control sellers but the difference is not significant. These two groups of sellers also have similar percentage of refunds and complaints, share of positive reviews as well as frequency of violating the platform's rules. The point estimates suggest that sellers with training access weakly out-perform control sellers for most of the quality metrics.

Focusing on sellers who actually participate in the training, we use the random assignment of the training access as an instrument for actual participation and an indicator for taking up any tasks during the sample period as the first stage variable. Column 1 of table A3 shows the first stage specification as equation 1. On average 25.7% of treated sellers took up some tasks. The rest of the table presents the two-stage least-square estimates on sellers' performance. For sellers taking up the tasks, they earn 6.6% higher revenues and attracted 5.2% more visitors to their sites. We do not find significant impacts on conversion rates. Comparing sellers who took up some tasks to those who did not in the treatment group, it is obvious that training participants significantly out-perform non-participants along all performance measures (table A4).

We analyze temporal variations of the treatment effects in A.1 and conclude that the temporal variations are small. The experimental design limits the baseline heterogeneity we could capture to basic types, locations and actions on the first day. We discuss the heterogeneous treatment effects in details in appendix A.2. In a nutshell, we do not find significant difference by sellers' registration type or level of preparedness[28]. Instead, offline business environment could play a role. Sellers coming from less-developed regions are less likely to participate in the training and the training is less useful for these sellers.

The magnitude of the treatment effect on revenues is positive but small. Over the nine-

[28]We measure level of preparedness by whether or not sellers post products on the first day of registration.

month period, treated sellers earned $1.8 million higher total revenues. Assuming the treatment effect is of similar magnitude for all cohorts of the new sellers, all the two million treated sellers combined could earn about $4.7 million higher revenues. Higher revenues could be a result of market expansion and business stealing. We discuss these two possibilities in section 4 and section 5.

Sellers' Strategies Next we discuss how the new sellers' strategies change behind the observed increase in revenues and the number of visitors. We focus on observable strategies including pricing, product offerings, marketing and customer service. We find treated sellers change their marketing strategies and slightly improve customer service quality, but do not behave differently on other dimensions.

Table 3 presents estimated coefficients β in specification 1 on the treatment indicator, where each cell represents to a separate regression on a specific outcome variable. Although a number of the training tasks focus on technical and administrative barriers sellers may face when setting up their online stores, the training has limited impacts on incentivizing market participation in terms of posting products or putting down deposits. The results suggest that closing the knowledge gap alone is not enough since many offline constraints are still limiting. For example, sellers need to find sources of supplies and have available funds to cover operation costs. Eventually 65.0% of sellers in the treatment and 64.8% of sellers in the control group posted products, but the difference is not significant. Moreover, treated sellers are not accelerating the speed of posting products among the subset of sellers who posted products after the first day of entry (table A2 column 1 and 2). Similarly, treated sellers are no more likely to put down security deposits (table 3). As expected, sellers do not behave differently in terms of number of products offered, likelihood of moving into different sectors or setting different prices. These strategies are not covered by the training and are more affected by offline environment.

Treated sellers are more likely to follow the platform's recommendations to adopt supplementary tools that help improve quality of customer service. Specifically, we find that treated sellers have slightly shorter average response time when consumers making inquiries and have higher conversion rates among consumers who made inquiries (table 3 section on customer service). These results are driven by treated sellers' higher likelihood to adopt the AI-backed customer assistant to help answer consumers' basic questions. Although we do not observe any difference on customer ratings on sellers' service quality, the improvement in the customer service quality could still contribute to treated new sellers' revenue increases.

Training helps improve treated sellers' marketing skills. Treated sellers have more products participating in pay-per-click ads where sellers bid for better search rankings with specific search keywords and they have higher share of visitors coming from the paid channels[29]. In addition to advertising, treated sellers are also more likely to participate in the limited-time promotional events that the platform regularly organizes[30]. The products on sales get preferential treatment in search rankings and additional exposure since consumers are able to find these product from other channels besides the main search and recommendation program.

Marketing is an indispensable part of online business operations and attracting visitors is the key to success. However, marketing capacity and quality might not perfectly correlate (Hu and Ma, 2020). The training intervention that either improves sellers' marketing skills or raise sellers' awareness of marketing helps new sellers accumulate more consumers[31]. As a result, incumbents and non-treated new sellers could have fewer visitors. Such reallocation has ambiguous implications on consumers. The ambiguity is in fact a common concern for typical training interventions that promote specific groups because of potential negative selection. In such cases, consumers might interact with lower quality firms more often due to the interventions. We assess the impacts of the training on consumers in section 4.

4 Effect of the Training on Consumers

As discussed in the previous section, the entrepreneur training causes treated new sellers to attract more consumers to their sites and improves customer service quality. As a result, consumers' experience on the platform could be affected when they interact with different types of sellers. In this section, we evaluate the impacts of the training on consumers by answering the following questions. First, when a consumer visit more treated or control new sellers during a search session, is she more likely to find what she needs and purchases from some sellers in the set? Second, when a consumer visit both new sellers and incumbents in a search session, from whom she is more likely to purchase? Changes in consumers' search experience could affect overall purchase probability (market expansion) and choices within a set of visited sellers (market allocation). Empirically, we use the detailed consumer-seller matched browsing data and exploit variations in the compositions of sellers that consumers visit given their interests and search efforts. The set of sellers a consumer visits in a search session is determined by platform's search algorithm and her own browsing behaviors. While both are affected by consumers' characteristics and past

[29]The typical paid channels include pay-for-clicks ads in search results, advertising spots in the AI-powered recommendations, headline figures and social media campaign.

[30]The classic sales events the platform regular offer are product specific limited time discounts. Sellers select the set of products to participate and submit applications to the platform in order to be included.

[31]Since we do not observe sellers' operating or marketing costs, we are unable to identify if sellers' investment in the marketing ended up yielding higher profits. However, the bottom line is that training should have no direct impacts on sellers' available financial resources to invest in marketing.

behaviors, the search algorithm has some random assignment procedures especially when matching new sellers[32]. For consumers, explicitly picking out new sellers from the search results is nearly impossible without actually visiting sellers' sites.

4.1 Training and Market Expansion

Sample Construction To evaluate how interacting with new sellers affect consumers' subsequent experience, we identify consumers with the same interests (searching the exact same query) and the same search efforts (visiting the exact same number of sellers) on the same day. These consumers ended up visiting different set of sellers. To be precise, we construct a consumer-search session sample using administrative data from the platform in the following steps. First, we draw a random sample of new sellers from the experimental sample[33]. Next, we identify a set of consumers who have visited the sampled sellers' sites between August 1 and December 31, 2019 and obtain the search keywords they used to find these sellers. Then, we find another set of consumers who searched the exact same keywords on the exact same day and visited the exact same number of sellers as the matched consumers did but only visited incumbents[34]. As mentioned, sellers that each consumer visited are not randomly selected because both search rankings and consumers' own browsing behaviors are endogenous. However, without platform's explicit promotion in the search outcomes, it is almost impossible for consumers to specifically look for new sellers when searching for products. Our empirical strategy therefore exploits variations in platform's search algorithm which determines the pool of sellers that consumers could access. For these two groups of consumers who visited some new sellers or only visited incumbents, we obtain their search and purchase history for a month before and a month after the event. We aggregate the final sample to consumer-search session level. A search session is a search query-search efforts-date combination. For each consumer-search session, our main outcome measures are whether or not the consumer actually places an order with any sellers in the set and her total spending. Each observation therefore cor-

[32]We do not have access to the actual search algorithm and the algorithm is too complex and too dynamic to be summarized as simple rules. Discussions with internal engineers suggest that the algorithm does have random component when matching for new products and new sellers but there are no explicitly rules on how such matching associates with consumers' characteristics.

[33]We have to take a sub-sample from the full sets mainly due to computational reason. All sellers in the experimental sample combined attract more than 10 million visitors on a typical day.

[34]We limit the number of matched consumer per search query-search effort-date set at 50 for computational reason. Such sampling implies that we over-sampled sellers with fewer visitors and keywords that are less popular. We also require that consumers need to visit at least three sellers because consumers who visit only one or two sellers might have different mindsets. On the one hand, those consumers might not be serious in purchasing because of their limited search efforts. On the other hand, those consumers might be looking for very specific sellers, especially those they have purchased from previously as the algorithm tends to promote these sellers. We are unable to separate these two possibilities and these motivations could result in opposite purchasing behaviors.

responds to a specific consideration set that the consumer make purchase decision from. We obtain sellers' pricing, number of products offered and ratings as the main control and average these measures to consumer-search session level. For the consumers, we obtain their search and purchase history around the time of the search session, which allows us to summarize consumers' characteristics and preferences.

The final sample consists of 1,381,273 consideration sets (consumer-search sessions) spanning 153 days in the second half of 2019. Table A8 summarizes the main variables. The complete sample consists of 515,748 consumers. On average, each consumer appears in 2.87 search sessions, where 44.2% of consumers appear in only one search session. These consumers searched 13,593 distinct keywords, spanning most of the popular sectors on the platform. In 18.1% of the search sessions, consumers placed an order from some sellers they visited on the same day. Average consumers visited 4.89 sellers per search session and for 61.1% of the sessions consumers visited three or four sellers.

Interacting with new sellers is rare. In our sample, consumers visited treated new sellers in 6.5% of search sessions and control new sellers in 9.2% of sessions[35]. It is evident that new sellers that appear in consumers' consideration sets are larger and more successful than average new sellers in the experimental sample, especially since a significant proportion of the new sellers do not have or have very few visitors. Therefore, the results on consumers do not speak to the average new sellers, but only new sellers in the top end of the distribution whom the consumers can actually encounter when searching on the platform.

The characteristics of consumers who visited new sellers are different from those who only visited incumbents. In these search sessions, consumers visit more sellers. Consumers also spend more money and search more intensively in the week prior to the search event. For search sessions involving new sellers, consumers visited three or four sellers in only 51% of these sessions (figure A4). Conditioning on search effect, keyword and date, consumers visiting some new sellers spend 38% more than consumers only visiting the incumbents (figure A5). We address the selection on consumers side in the empirical analysis with consumer and search session fixed effects as well as a rich set of controls for consumers' characteristics at the time.

Empirical Strategy When a consumer visits more treated or control new sellers in a search session, is she more likely to find what she needs and make purchase? We answer the question using the consumer-search session sample with the following specification:

$$Y_{is} = \beta^t \mathbf{T}_{is} + \beta^c \mathbf{C}_{is} + \mathbb{X}_{is}\gamma + \alpha_s + \alpha_i + \epsilon_{is} \tag{2}$$

[35] As described in the sample construct, the probability of interacting with new sellers are calculated based on the sample and may not be the same as the probabilities on the entire platform. We over-sampled sellers attracting fewer visitors and less popular keywords. The actual probability of having some new sellers in the consideration sets over all the search sessions should be lower than what we have here in this sample.

for consumer i and search session s. The main variables of interests are $\mathbf{T}_{is}$ and $\mathbf{C}_{is}$, defined as share of treated or control sellers in the consideration set for consumer i in search session s. Since neither the platform nor consumers know if new sellers have access to the training, the comparison between treatment and control new sellers is less subject to selection. We add the fixed effect α_s to capture variations across search sessions. Controlling for the search query addresses heterogeneous responses when consumers searching for different types of products. For example, consumers tend to explore more options when they search for horizontally differentiated products such as clothing than when they search for vertically differentiated products such as home supplies. Controlling for number of sellers that consumers visited in a search session helps alleviate two competing concerns. On the one hand, consumers visiting more sellers are more dedicated as they invest more search efforts, which increase their likelihood to purchase. On the other hand, consumers browsing more sellers might also be less satisfied with their previous matches, which forces to them to search more intensively but could also lower their overall probability to purchase. Therefore, the size of the consideration set serves as a proxy for intensity and quality of the searches. Similarly, we control for date because the platform organizes many promotional events year round and these events could have differential impacts on new sellers and incumbents. We control for consumer time-invariant characteristics with α_i to address consumers' idiosyncratic variations in their pickiness, purchasing power, experience and familiarity with the platform. In addition, we also add control variables $\mathbf{X}_{is}$ that includes sellers' average pricing level, number of products offered and ratings in the consideration sets[36] as well as consumer i's total spending and search intensity in the previous week. This specification does not control for consumers' search query specific preferences, e.g. a consumer might be unusually picky when choosing printing paper even though her fellow shoppers view printing paper a homogeneous product. Such idiosyncratic preferences could affect consumers' search behaviors but it is unclear how would such taste difference bias the way consumers interact with new sellers.

Results Table 4 presents the results of the specification 2 on the purchase probability and log order size. On the day of visit, compared to visiting a set of sellers with incumbents only, a consumer is 11% more likely to make a purchase if her consideration set only consists of new sellers (column 1). Conditioning on size of the consideration set, consumers are 1.9% more likely to make a purchase with one more treated or control new seller in the set (table A10 column 1). Incorporating the possibility that consumers might place an order in later days, we reach similar conclusion that having more new sellers in the consideration sets significantly increases the likelihood for consumers to make purchases and the magnitude of such increase in the purchase likelihood is substantial. Having a search set

[36]Ratings are determined by cumulative number of positive reviews sellers get, hence they reflect sellers' size more than quality. Price level are measured in log.

with new sellers only increases total spending by at least 6.5% (column 4 in table 4), while having one extra new sellers in the consideration increases total spending for the specific search session by 1.1% (column 4 in table A10). Higher purchase probability indicates that consumers benefit from better matching quality in the searches. Higher intensity of interactions with new sellers therefore leads to market expansion.

Comparing control new sellers to treated new sellers who appear in consumers' consideration sets, we find that for the most part interacting with these two types of new sellers lead to similar increase on consumers' purchase probability. We test the difference in the estimated coefficients β^c and β^t and find that the difference between these estimated coefficients is small and insignificant expect for log total spending (column 4 table 4). If we use number of treated new sellers as the main explanatory variable rather than share of the new sellers in the consideration sets, there is no significant difference between estimated β^c and β^t (table A10). Since the training increases treated new sellers' likelihood to appear in the consideration sets and also improves their customer service quality, there is both selection and treatment effect. Overall we do not find negative selection because of the training, since interacting with treated new sellers is associated with higher matching quality similar to that of visiting non-treated new sellers. More importantly, consumers' search efficiency rises as they interact with both types of new sellers compared to having only incumbents in the consideration sets. The gap in purchase probability implies that there could be significant frictions that hinder new sellers' growth.

4.2 Training and Market Reallocation

In this section, we ask, when a consumer visits both new sellers and incumbents in a search session, from whom she is more likely to purchase? We construct a consumer-seller matched pair sample to analyze consumers' choices and the resulted market allocation.

Data and Sample Construction To construct the sample, we restrict the attention to consideration sets where consumers have visited at least one new seller and have made purchases from some sellers in the sets. We construct a consumer-seller pair level sample where each pair is associated with a specific consideration set hence a search session defined by search query - search effort - date combination as before. For each consumer-seller pair, we use whether or not a consumer makes a purchase from the specific seller within a given period of time as well as the size of the order as the main outcomes. In the cases when a consumer makes a purchase, we obtain quality measures on whether or not the consumer requests returns or refunds as well as whether or not she makes repeat purchases from the same seller in the following month. To control for sellers' strategies and characteristics, we again collect data on sellers' pricing level, number of products offered and ratings. We restrict the sample to the set of consumers who appear in at least two search sessions.

The final sample consists of 300,273 consumer-seller pairs belong to 42,004 consumers' 61,280 consideration sets (defined per search session - consumer) spanning 153 days. Table A9 presents summary statistics of main variables used. These search sessions are the same as those we analyzed in section 4.1 except in this sample we obtain more detailed consumer-seller level interactions associated with these sellers. There are 98,631 sellers appearing in the sample, with 3,687 sellers belonging to the treatment group and 7,329 sellers in the control group. As discussed earlier, these new sellers are highly selected as they are much larger and more active than average new sellers. Average sellers have 18.8% chances to be selected by consumers on same day and 21.9% of sellers get orders within a week. Average order size is \$31.5 and the median order size is \$20.4. For 9.1% of purchases consumers request refunds or returns and consumers place repeat order in the following month in 4.9% of the cases. We use quantity weighted average prices at seller level as proxy for seller's pricing level[37].

Empirical Strategy To test the impacts on allocation, we use the following specification:

$$Y_{ijs} = \beta^t \mathbf{T}_j + \beta^c \mathbf{C}_j + \mathbb{X}_j \gamma + \alpha_{is} + \epsilon_{ijs} \tag{3}$$

The specification includes consideration set α_{is} fixed effects to address consideration set level heterogeneity. Consideration set is a set of sellers a consumer eventually chooses from in a search session. In this way we control for consumer - search session specific idiosyncratic variations and only evaluate consumers' choices between sellers in the sets. The outcomes of interests are $\mathbf{T}_j$ and $\mathbf{C}_j$, indicators for whether or not seller j belongs to treatment or control group. To test the differential impacts of interacting with treated and control new sellers, we compare the coefficients β^t and β^c. As before, we include a set of seller level controls $\mathbb{X}_j$ on seller j's pricing, number of products offered and ratings.

Results The estimated results are summarized in table 5. By restricting to consideration sets that consumers purchased from, we control for the possibility of market expansion and evaluate the market allocation between new sellers and incumbents. Consumers are more likely to choose new sellers, especially new sellers with access to training, over incumbents that appear in their consideration sets. Specifically, on the day of visit, consumers

[37] Almost all the sellers on the platform offer multiple products and in the search channel consumer access a particular seller from the product page. However, due to data limitation and complexity of the pricing strategies, we do not observe real-time prices that the consumers observe. Moreover, since we aggregate the outcomes to seller level and consumers could browse and purchase multiple products from the sellers they visit, it is unclear how to aggregate prices without observing what products the consumers visit. Therefore, we use seller level, quantity weighted pricing as a proxy for the seller's pricing level. Current algorithm encourage sellers to design their pricing strategies so that they could target a specific group of consumers with comparable purchasing power. Therefore, the variations in prices across products in a store could be limited than across seller variations.

are 5.9% more likely to choose a treated new seller than an incumbent in the same consideration set. Comparing treated new sellers with control new sellers, consumers' purchase probability with the former is significantly higher by about 4.2%. The differential impacts of interacting with treated or control new sellers are similar if we use purchase within a week or amount spending as the outcomes. In column 2, 4 and 6 we include seller level controls. The coefficients are quantitatively similar. Results here suggest that consumers' higher purchase probability with new sellers is not driven by sellers' charging different prices.

Using the consumer-seller matched sample we show that training benefits the consumers by improving their matching quality and confirm that the improvements on matching is indeed generated by consumers interacting with new sellers. The results shut down the potential negative selection induced by the training. As the training increases treated new sellers' likelihood to appear in consumers' consideration sets, the reallocation could generate market expansion and market reallocation as a result. Such shift in allocation of consumers' visit likelihood could benefit the consumers but may come at the costs of control new sellers and incumbents. Current evidence suggest that new sellers have higher quality than incumbents so that reallocation improves overall efficiency. We quantify and decompose such impacts in section 5.

Besides purchase probability, table 6 shows how post-purchase experience may differ when consumers purchase from new sellers versus the incumbents. We restrict the sample to consumer-seller pairs where consumers make purchase on the day of visit and use specification 3 to evaluate the consequences on the likelihood of return, refund and repeat purchase[38]. Overall, placing an order with new sellers does not have significant negative impacts on consumers' likelihood to request return or refunds (column 1 and 2). Moreover, there is no difference in the likelihood of making repeat purchase from new sellers versus incumbents (column 3). These results show that while consumers are more likely to purchase from new sellers, their post-purchase experience after purchasing from the new sellers are no worse than their experience after purchasing from incumbents. Hence, higher purchase probability does not come at a cost of lowering purchase quality.

5 Quantify and Decompose the Impacts of the Training

Motivated by reduced-form evidence, we use a structural model to characterize growth barriers new sellers face. Our model focuses on consumers' purchase decisions given their consideration sets and use variations in consumers' choices to identify the main source of frictions: the mismatch between sellers' true quality and total number of visitors they

[38]We could not add seller level controls because the model cannot be identified with due to large number of fixed effects and control variables. Hence the control variables only include consumer baseline characteristics.

manage to obtain. To close the model, we use a flexible function to characterize platform's matching rule based on observable sellers' performance, namely their lagged number of visitors and conversion rates[39].

To evaluate the welfare implications of the training, our model emphasizes how the training changes sellers' likelihood to appear in consumers' consideration sets that consequently changes matching quality and welfare. In the absence of the training intervention,more incumbents and control new sellers will enter consumers' consideration sets. Because of the training, treated sellers adjust their strategies accordingly to attract more visitors to their sites. Treated new sellers capture more attention from the consumers imply that the incumbents (and control new sellers) will be less likely to enter consumers' consideration sets. That is, training induces change the composition of sellers in consumers' consideration sets. We expect the negative spillover on the non-treated new sellers to be limited because vast majority of sellers on the platform are incumbents, so that if reallocation occurs randomly among all sellers then by chance most of the market reallocation will come from incumbents. In the model, we take consumers' searching and browsing behaviors as given and only consider the market reallocation among different types of sellers[40].

5.1 Model Setup

Consumer Demand As mentioned, we do not explicitly model consumers' search process and how they arrive at the observed consideration sets. Instead, we take these consideration sets as given and consider the conditional purchase decisions. Specifically, a consumer $i \in I$ searches a query and generates a consideration set K_i. Each set K_i consists of a group of sellers $j \in K_i$ that could be either (treated or control) new sellers or incumbents. Consideration sets K_i could have different sizes which we do not model. Consumer i solves the following maximization problem to choose from which seller $j \in K_i$ she wants

[39]We do not directly model supply side responses because empirically changes in strategies that could affect consumer demand such as price adjustments and product introductions are in fact rare. Instead, most of the actions sellers take concentrate on marketing, which is captured by number of consumers they attracted in previous periods.

[40]On the one side, treated new sellers are more likely to appear in consumers' search sets, resulting market reallocation from incumbents control new sellers to treated new sellers. By closing the knowledge gap, training helps to ensure that new sellers could participate in the competition for consumers' attention. On the other hand, consumers may change their search behaviors in response to changing composition of the sellers they visit. Consumers need to spend less efforts to search if quality of the matches improve, which is welfare improving for the consumers but could limit fellow sellers' chances to appear in consumers' consideration sets. Alternatively, matching with higher quality sellers in the search sessions may induce consumers to do more searches in the future, raising the likelihood of purchase from other sellers. This channel could potentially benefit all the sellers as market for consumers' attention expanded. Comparing to market reallocation and direct impacts of changes in consideration sets on purchase, changes in consumers' search behaviors are second-order.

to purchase:

$$\max_{j \in K_i} U_{ij} = V_{ij} + \epsilon_{ij} = x_j \beta - \alpha p_j + \xi_j + \epsilon_{ij} \tag{4}$$

p_j is price level seller j charges and x_j is the set of strategies and characteristics seller j adopts that might affect consumers' purchase decisions. ξ_j is the unobserved seller j's underlying quality and is our main object of interest. Sellers with higher quality ξ_j yield higher utility for all consumers visiting their sites. The main source of friction the model captures comes from the mismatch between sellers' quality ξ_j and their likelihood to appear in consumers' consideration sets. We do not explicitly model treatment effect but since we use post-treatment data to estimate ξ_j, ξ_j captures both sellers' underlying quality and effect of the treatment. ϵ_{ij} is the I.I.D. consumer-seller idiosyncratic preference that reflects unobservable components affecting consumers' decisions.

Assume that ϵ_{ij} follows type I extreme value distribution and normalize consumers' outside option of not purchasing from any seller to have zero utility, we get the following familiar logit formulation

$$P_{ij} = \frac{\exp(x_j \beta - \alpha p_j + \xi_j)}{1 + \sum_{k \in K_i} \exp(x_k \beta - \alpha p_k + \xi_k)}$$

where P_{ij} is consumer i probability of purchasing from seller j. We later enrich the baseline model by adding consumer side heterogeneity, sector specific fixed effect ξ_s and sector specific price coefficient α_s.

Endogenous Strategies One major concern with the baseline model is that pricing level p_j and strategies such as number of products offered in x_j could correlate with ϵ_{ij}, which bias the estimated coefficients. To address such concern, we use a set of instruments to jointly determine number of products posted prod_{jt} and pricing level p_{jt} with

$$\begin{bmatrix} p_{jt} \\ \text{prod}_{jt} \end{bmatrix} = \mathbf{Z}\beta^{fs} + \xi_j^{fs} + \xi_t^{fs} + \xi_s^{fs} + \epsilon_{jt} \tag{5}$$

ξ_j^{fs}, ξ_t^{fs} and ξ_s^{fs} are seller, time and sector fixed effect. The set of instruments Z are variables that capture the stringency of platform's rule enforcement. These instruments include number of visitors weighted average frequency of different types of rule violations as well as shares of sellers identified as frequent rule violators under various standards in sellers' affiliated sectors. The most common rule violations include infringing intellectual property rights, selling counterfeits and providing false or misleading product information. The platform enforces comprehensive rules not only to ensure that sellers obey relevant state regulations but also to maintain the well-functioning of the market. In the cases of rule violations, the platform could downgrade sellers in the search rankings, remove access to sellers' products or even their sites and in some cases call for legal solutions. Such punishment could have significant impacts on sellers' business operations. The platform frequently adjusts the design and enforcement of the regulations as business environment fluctuates. We exploit changes in the strictness of rule enforcement at sector-month level. When the platform strengthens the rule enforcement, sellers could be more cautious about posting more products, charging extreme prices or engaging in unruly promotions. On the other side, when the platform enforces stricter rules, rule-obeying sellers could benefit as the platform regulates the behaviors of their unscrupulous competitors, allowing the rule-obeying sellers to increase their market shares.

Matching The rule for matching sellers and consumers is the most important device that the platform has to improve consumers' experience and to support promising sellers. We simplify the complex matching rules used in the search and recommendation algorithm by highlighting the reliance on previous period's conversion rates and number of visitors. In the matching, conversion rate directly reflects seller-specific consumer demands that are affected by sellers' underlying quality ξ_j. Last period's total number of visitors summarizes sellers' characteristics especially their marketing skills and the impacts of the training on attracting consumers. Sellers' strategies including their participation in the training do not directly factor in the matching process, only the lagged results are. We capture the evolution of the number of visitors over time with the following model:

$$\mathcal{T}_{jt} = f(\mathcal{T}_{jt-1}, C_{jt-1}, C_{jt-2}) \tag{6}$$

$\mathcal{T}_{jt}$ is current period total number of visitors for seller j and $\mathcal{T}_{jt-1}$ is previous period's number of visitors[41]. C_{jt-1} and C_{jt-2} are conversion rates in the previous two periods.

5.2 Estimation

We estimated the model using simulated maximum likelihood following Train (2009). To better fit the empirical setting, we make the following changes to the basic model.

Consumer Demand We use the consumer-seller pair sample to estimate the demand parameters in particular sellers' ξ_j. Appendix B.1 describe the detailed sample construction process. Due to computation constraints, we sample a subset of sellers for actual estimation. The final sample is a seller-consumer matched pair dataset and we explore seller level variations. x_j includes number of products sellers offer as well as sellers' ratings and p_j is average price level that sellers charge[42]. To account for sector level heterogeneity, we

[41]We again convert number of visitors to log scale to improve the fit because of the skewness of the distribution.

[42]As mentioned before, ratings capture sellers' size more than quality. The price level seller charges is the quantity weighted prices of all products sold by the sellers during the day. Such weighted price reflect relative

add sector-specific intercept ξ_s in the baseline model and enrich the baseline model by estimating sector specific price coefficient α_s.

We jointly estimate consumer demand and sellers' strategy using the instruments described above. The final set of instruments include number of visitors weighted average frequency of rule violations and share of sellers labeled as frequent rule violators in sector s over the previous 30-day period where we define time as the time when actual visits occur. In the baseline model we do not explicitly account for temporal variations. More detailed estimation procedure are described in B.2.

Matching To estimate the matching rule, we use the the new seller panel. We use the following specification to specifically distinguish new entrants with no previous history and sellers with zero conversion rates in the previous periods from the rest:

$$T_{jts} = f(T_{j,t-1}, C_{j,t-k}) + g(I_{j,t-k}(t-k=0)) + h(I^c_{j,t-k}(C_{j,t-k}=0)) + \xi^{trf}_k + \xi^{trf}_t + e_{jts} \quad (7)$$

$C_{j,t-k}$ is lagged conversion rates in past periods $k = 1, 2$, $I_{jt-k}(t-k=0)$ is an indicator for the initial two periods[43] and $I^c_{j,t-k}(C_{j,t-k} = 0)$ are indicators for lagged conversion rates being exactly zero. The specification also includes product category, calendar time and relative month fixed effect. In the baseline specification we start with linear functions for $f(\cdot)$, $g(\cdot)$ and $h(\cdot)$. In this setup, all the right-hand-side variables are determined in the previous periods.

5.3 Estimation Results

Following the procedure described in section 5.2, we estimate the baseline model and quantify the welfare implications of the training program with counterfactual exercises.

Panel A of figure A6 plots the distribution of estimated sector fixed effect ξ_s. Panel B of figure A6 plots the distribution of price elasticity for sellers in the sample. The average price elasticity is -0.22. The elasticity we estimate here is much smaller than typical elasticity observed in the literature (Broda and Weinstein, 2006). The difference occurs because consumers are choosing between products in their consideration sets, rather than choosing among all the products offered in the market. When constructing the consideration sets, consumers already restrict their choices to a narrower range of prices. In our sample, average standard deviation of prices among all sellers that some consumers visited is 3.14 times higher than average standard deviation of prices among sellers in the consumer-specific consideration sets. Therefore estimated price elasticity with the consideration set is lower.

popularity of products sellers offer as well as any sales or promotions sellers offer.

[43]Sellers have no past history on number of visitors and conversion rate other than characteristics in the initial periods hence are subject to different matching rule.

The main parameters of interests in are ξ_j for $j \in J$. ξ_j captures seller's post-treatment underlying quality, incorporating selection and treatment effect. The final sample contains 52,241 sellers out of which 8.33% are new sellers. 28% of sellers in the sample have some purchase records. Table 7 presents the distribution of ξ_j for different subset of sellers. On average, new sellers have higher ξ_j than incumbents, both among sellers with and without purchase records. The difference is summarized in table 8. Among sellers with purchase records, estimated quality of new sellers is 17% higher than the incumbents and among those without purchase records, new sellers' estimated quality ξ_j is 11.2% higher. As shown in the distribution of ξ_j (figure 4), the difference is not driven by a small set of extremely high quality new sellers whose ξ_j land on the right tail of the distribution. Instead, the results are driven by median new sellers having higher underlying quality than the median incumbents. The sample of new sellers is a highly selected subset from all new sellers. These results confirm what we found in the reduced-form analysis: new sellers have higher underlying quality, allowing them to out-compete incumbents in the same consideration sets. To unpack the welfare implications of the training intervention, we turn to counterfactual scenarios in the matching.

5.4 Counterfactual: Welfare of the Training

We analyze the welfare consequence of the training by considering how changes in the likelihood for different types of sellers to appear in consumers' consideration sets induced by the training affect consumer surplus and sellers' revenues.

To conduct the counterfactual analysis, we randomly sample a subset of sellers along with their associated search sessions. The potential pool of sellers a consumer searching a particular keyword could choose from consists of all sellers visited by some consumers searching that keyword in the full sample. For each seller, we obtain their strategies x_j and estimated quality ξ_j. Based on the estimated results, we calculate consumers' utility when they visit a particular seller j, V_j. In the baseline model, V_j is the same across consumers. To construct the consideration sets, we hold constant the consumer-search session pairs and sample sellers from pool of sellers associated with the search keyword according to certain sampling weights. We match the number of sellers sampled with what we observe in the data. Therefore in the counterfactual, we only vary composition of the consideration sets and hold everything else constant. Since training increases treated new sellers' likelihood to appear in consumers' consideration sets, we evaluate the welfare of the training by restricting training participants' probability of being sampled, as described below. More details about construction of the counterfactual consideration sets are described in B.3.

Baseline: Predicted Number of Visitors In the baseline version, sellers' sampling weights is given by their predicted numbers of visitors as determined by the empirically estimated matching rule described in section 5.3. The current logit-specification allows us to calculate consumer surplus as

$$CS = \frac{1}{\alpha}\sum_i \log[\sum_{j\in K_i} \exp(V_j)]$$

Sellers' revenues are given by the probability of being chosen and price level they charge:

$$R_j = \sum_i \frac{\exp(V_j)}{1+\sum_{k\in K_i}\exp(V_k)} p_j$$

Sellers will not earn any revenues if they do not appear in the consideration sets. Table 9 summarizes the results of welfare decomposition. In the baseline, new sellers only capture 6% of total revenues even though they represent 8.3% of sellers in the pool.

Welfare Impact of the Training Since the training intervention increases treated new sellers' likelihood to appear in consumers' consideration sets, we estimate the impacts of the training by limiting some new sellers' appearance in the search results. Without the training, treated new sellers' likelihood to appear in the sampling pool should be similar to that of the control sellers. Assuming that only training participants are affected by the training, we evaluate the welfare of the training by randomly removing a subset of training participants from the sampling pool that consumers could choose from. As a result, more consumers will end up visiting non-participants, control new sellers and incumbents instead of training participants, which is what we expect in the absence of the training. Restricting the effect of the training on training participants allow us to account for selection into training participation. The average difference between the estimated quality of training participants and that of non-participants is small. Appendix B.3 discuss the details of the counterfactual exercise.

As presented in the first row of table 9, treated new sellers' revenue share drops by 7.7% in the absence of training. Total sellers' revenues decrease by 0.05% and consumer surplus decreases by 0.07% as a result of lowering the likelihood to visit higher quality training participants. Even though the magnitude of welfare loss is small in percentage term, the absolute magnitude of welfare loss is substantial because of the volume of total transactions facilitated by the platform. To decompose the source of revenue growth induced by the training, we compare the changes in the market shares by different types of consideration sets and sellers in these sets. Comparing to the baseline, revenues generated from consideration sets with treated new sellers drop significantly as we limited treated new sellers' presence. The drop suggests that most of the revenue growth induced by the training is driven by market expansion as consumers are more likely to purchase when treated new sellers appear in the consideration sets. For market reallocation, we compare the market share of different types of sellers if they appear in the consideration sets that have at least one new seller. The market share of the treated new sellers drop in these sets and control new sellers gain as a result.

Welfare Impacts of Scaling Up the Training To evaluate the potential welfare impacts if we scale up the training to provide access to all the new sellers, we evaluate the changes in welfare if we limit new sellers' presence in the sampling pool. We randomly remove some new sellers from consumers' sampling pool to match the reduced form estimate where we find that training increases new sellers' number of visitors by 9.7%. The third row of table 9 presents the results. In this case, total consumer surplus drops by 0.08% and total sellers' revenues decreases by 0.3%. Both types of new sellers lose their market shares by over 8% as expected. The drop in revenues is mostly driven by the fact that as fewer high quality new sellers appear in consumer's consideration sets, consumers make fewer purchases overall. The effects of market reallocation among different types of sellers in these sets are small. Scaling up the training would lead to greater increases in sellers' revenues than in consumers' surplus, which benefits the platform even more in the short run.

6 Conclusions

In this paper we study how a business training intervention can be an effective way to lift growth barriers new entrants face in a competitive e-commerce platform where sellers face demand-side frictions. Leveraging the experimentally randomized access to the training and the unique consumer-seller matched searching and browsing data, we show that the business training helps new sellers increase their presence in consumers' consideration sets and earn higher revenues. The resulting changes in the composition of consumers' consideration sets are beneficial for the consumers as they enjoy higher matching efficiency without lowering the quality of the purchase when more treated new sellers enter their consideration sets. Using a structural model where we highlight the mismatch between number of consumers sellers acquire and their quality, we show that the training increases consumers' welfare and total revenues by limiting the extent of misallocation. The improved matching quality as well as new sellers' higher service quality could also lower consumers' search costs, which we do not account for in the current analysis. Enhancing matching quality to improve consumers' experience and promising sellers' growth potential is consistent with the platform's long-term profit-maximizing goal. In the short-run, market expansion and sellers' increased engagement with online marketing also contributes to the platform's profits. Overall, the platform occupies an unique position to implement such interventions to support promising new sellers with active engagement.

Our findings provide one of the first sets of direct empirical evidence on the welfare implications of an intervention that supports subset of firms on the consumers. The fact that consumers do not experience negative selection in this context is because the training reduces market frictions and improved matching quality for the higher quality new sellers. As the market operators, the platforms could play critical roles in lifting growth barriers with proper interventions and doing so is in alignment with the platforms' incentives as profit maximizing firms in both short and long run. Although large e-commerce platforms' looming market power should not be overlooked, these platforms indeed create enormous opportunities for SMEs and they have incentives as well as capacity to take more active roles to foster the efficiency and equality in online markets they host.

References

Anderson, Stephen J., Rajesh Chandy, and Bilal Zia, "Pathways to Profits: The Impact of Marketing vs. Finance Skills on Business Performance," 2018, *64* (12), 5559–5583.

Arkolakis, Costas, Theodore Papageorgiou, and Olga A Timoshenko, "Firm learning and growth," *Review of Economic Dynamics*, 2018, *27*, 146–168.

Atkin, David, Amit K Khandelwal, and Adam Osman, "Exporting and firm performance: Evidence from a randomized experiment," *The Quarterly Journal of Economics*, 2017, *132* (2), 551–615.

Bai, Jie, Maggie Chen, Jin Liu, and Daniel Yi Xu, "Search and Information Frictions on Global E-Commerce Platforms: Evidence from AliExpress," 2020.

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman, "Six randomized evaluations of microcredit: Introduction and further steps," *American Economic Journal: Applied Economics*, 2015, *7* (1), 1–21.

Banerjee, Abhijit V and Esther Duflo, "Do firms want to borrow more? Testing credit constraints using a directed lending program," *Review of Economic Studies*, 2014, *81* (2), 572–607.

Berman, Nicolas, Vincent Rebeyrol, and Vincent Vicard, "Demand learning and firm dynamics: evidence from exporters," *Review of Economics and Statistics*, 2019, *101* (1), 91–106.

Bloom, Nicholas and John Van Reenen, "Measuring and explaining management practices across firms and countries," *The Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

__, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, "Does management matter? Evidence from India," *The Quarterly Journal of Economics*, 2013, *128* (1), 1–51.

Broda, Christian and David E Weinstein, "Globalization and the Gains from Variety," *The Quarterly journal of economics*, 2006, *121* (2), 541–585.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar, "The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in Mexico," *Journal of Political Economy*, 2018, *126* (2), 635–687.

Brynjolfsson, Erik, Yu Hu, and Michael D Smith, "Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers," *Management Science*, 2003, *49* (11), 1580–1596.

Calderon, Gabriela, Jesse M Cunha, and Giacomo De Giorgi, "Business Literacy and Development: Evidence from a Randomized Controlled Trial in Rural Mexico," *Economic Development and Cultural Change*, 2020, *68* (2), 507–540.

Chen, Maggie X and Min Wu, "The value of reputation in trade: Evidence from Alibaba," *Review of Economics and Statistics*, 2020, pp. 1–45.

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe, "Using big data to estimate consumer surplus: The case of uber," Technical Report, National Bureau of Economic Research 2016.

Couture, Victor, Benjamin Faber, Yizhen Gu, and Lizhi Liu, "Connecting the countryside via E-commerce: evidence from China," Technical Report, National Bureau of Economic Research 2018.

Cui, Ruomeng, Jun Li, and Dennis J Zhang, "Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on Airbnb," *Management Science*, 2020, *66* (3), 1071–1094.

Dinerstein, Michael, Liran Einav, Jonathan Levin, and Neel Sundaresan, "Consumer price search and platform design in internet commerce," *American Economic Review*, 2018, *108* (7), 1820–59.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar, "Keeping it simple: Financial literacy and rules of thumb," *American Economic Journal: Applied Economics*, 2014, *6* (2), 1–31.

Einav, Liran, Peter J Klenow, Benjamin Klopack, Jonathan D Levin, Larry Levin, and Wayne Best, "Assessing the gains from e-commerce," *Unpublished manuscript*, 2017.

Ellison, Glenn and Sara Fisher Ellison, "Match quality, search, and the Internet market for used books," Technical Report, National Bureau of Economic Research 2018.

Farronato, Chiara and Andrey Fradkin, "The welfare effects of peer entry in the accommodation market: The case of airbnb," Technical Report, National Bureau of Economic Research 2018.

Fitzgerald, Doireann, Stefanie Haller, and Yaniv Yedid-Levi, "How exporters grow," Technical Report, National Bureau of Economic Research 2016.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "The slow growth of new plants: Learning about demand?," *Economica*, 2016, *83* (329), 91–129.

Fradkin, Andrey, "Search frictions and the design of online marketplaces," *Working Paper*, 2015.

Goldfarb, Avi and Catherine Tucker, "Digital economics," *Journal of Economic Literature*, 2019, *57* (1), 3–43.

Gourio, Francois and Leena Rudanko, "Customer capital," *Review of Economic Studies*, 2014, *81* (3), 1102–1136.

Hjort, J and J Poulsen, "The arrival of fast internet and skilled job creation in Africa," *American Economic Review*, 2018.

Horton, John Joseph, "Misdirected search effort in a matching market: causes, consequences and a partial solution," in "Proceedings of the fifteenth ACM conference on Economics and computation" 2014, pp. 357–357.

Hu, Allen and Song Ma, "Persuading Investors: A Video-Based Study," Technical Report, Working paper 2020.

Hui, Xiang, Maryam Saeedi, Zeqian Shen, and Neel Sundaresan, "Reputation and Regulations: Evidence from eBay," *Management Science*, 2016, *62* (12), 3604–3616.

Jo, Yoon J, Misaki Matsumura, and David E Weinstein, "The Impact of E-Commerce on Relative Prices and Consumer Welfare," Technical Report, National Bureau of Economic Research 2019.

Lehdonvirta, Vili, Otto Kässi, Isis Hjorth, Helena Barnard, and Mark Graham, "The global platform economy: A new offshoring institution enabling emerging-economy microproviders," *Journal of Management*, 2019, *45* (2), 567–599.

Lipsman, Andrew, "Global Ecommerce 2019: Ecommerce Continues Strong Gains Amid Global Economic Uncertainty. eMarketer.," Technical Report, https://www.emarketer.com/content/global-ecommerce-2019 2019.

McKenzie, David and Christopher Woodruff, "What are we learning from business training and entrepreneurship evaluations around the developing world?," *The World Bank Research Observer*, 2014, *29* (1), 48–82.

__ **and Susana Puerto**, "Growing Markets through Business Training for Female Entrepreneurs: A Market-Level Randomized Experiment in Kenya," *American Economic Journal: Applied Economics*, forthcoming.

Muralidharan, Karthik and Paul Niehaus, "Experimentation at scale," *Journal of Economic Perspectives*, 2017, *31* (4), 103–24.

of Commerce, PRC Ministry, "E-commerce in China," 2020.

Piveteau, Paul, "An empirical dynamic model of trade with consumer accumulation," Technical Report, Working Paper 2016.

Rotemberg, Martin, "Equilibrium effects of firm subsidies," *American Economic Review*, 2019, *109* (10), 3475–3513.

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

Train, Kenneth E, *Discrete choice methods with simulation*, Cambridge University Press, 2009.

Ursu, Raluca M, "The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions," *Marketing Science*, 2018, *37* (4), 530–552.

World-Bank, *World Development Report 2016: Digital Dividends*, World Bank Publications, 2016.

Zhang, Dennis J, Hengchen Dai, Lingxiu Dong, Qian Wu, Lifan Guo, and Xiaofei Liu, "The value of pop-up stores on retailing platforms: Evidence from a field experiment with alibaba," *Management Science*, 2019, *65* (11), 5142–5151.

Table 1: Baseline Summary Statistics: Treatment Assignment and Participation

	Full Sample			Sellers in Treatment Group		
	Treatment	Control	(1) - (2) Difference	Participants	Non Participants	(4) - (5) Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Is Firm	0.268	0.264	0.004***	0.275	0.266	0.008***
	(0.443)	(0.441)	(3.33)	(0.446)	(0.442)	(3.18)
Female Owner	0.455	0.453	0.002	0.442	0.458	-0.16***
(among individual sellers)	(0.497)	(0.497)	(1.27)	(0.497)	(0.498)	(4.61)
Region: Coastal South	0.435	0.434	0.001	0.518	0.413	0.105***
	(0.496)	(0.496)	(0.28)	(0.500)	(0.492)	(3.573)
Region: West	0.118	0.118	0.0001	0.089	0.126	-0.038***
	(0.323)	(0.323)	(0.34)	(0.284)	(0.332)	(21.41)
List Products on Day One	0.213	0.212	0.0005	0.245	0.204	0.04***
	(0.409)	(0.409)	(0.44)	(0.430)	(0.403)	(16.04)
Number of Listed Products	1.539	1.535	0.004	2.307	1.342	0.965***
	(2.166)	(2.164)	(0.65)	(1.988)	(2.167)	(80.81)
Traffic	2.813	2.807	0.006	4.028	2.501	1.527 ***
	(2.289)	(2.283)	(0.91)	(2.259)	(2.191)	(115.35)
Conversion Rate	0.051	0.051	0.0002	0.055	0.050	0.006***
	(0.161)	(0.162)	(0.52)	(0.120)	(0.171)	(7.34)
Revenues	2.145	2.134	0.01	3.842	1.708	2.134***
	(3.312)	(3.304)	(1.144)	(3.726)	(3.048)	(100.66)
Observations	177,026	535,092	712,118	36,189	140,837	177,026

Notes: Columns 1, 2, 4, and 5 present means and standard deviations (in parentheses). Columns 3 and 6 show the difference in means across the treatment and control group (training participants and non-participants) in the full sample with the corresponding t-statistics in parentheses. Sample for female dummy further restricted to sellers that are not registered as firms. Participation is defined as having taken up any tasks during the nine-month period. Traffic, conversion rate, revenues and number of product posted are for the first month outcomes. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 2: Overall Treatment Effects on Sellers' Performance

	Dependent variable:							
	Log Revenues	Any Revenues	Revenues	Log # Visitors		Log # Buyers		Conversion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.017***	0.002**	0.026**	0.013**	0.024***	0.008***	0.018***	0.0001
	(0.006)	(0.001)	(0.010)	(0.006)	(0.008)	(0.003)	(0.006)	(0.0002)
Dep Var Mean	1.39	0.19	7.14	1.73	4.29	0.57	1.4	0.04
Sample	Full	Full	Earn Revenues	Full	Have Visitors	Full	Have Visitors	Have Visitors
Observations	6,409,062	6,409,062	1,253,284	6,409,062	2,593,762	6,409,062	2,593,762	2,593,762
R^2	0.132	0.152	0.081	0.207	0.111	0.105	0.070	0.043
Adjusted R^2	0.132	0.152	0.081	0.207	0.111	0.105	0.070	0.043

Notes: Dependent variables are monthly outcome for all sellers in the new seller sample. Traffic (number of visitors), number of buyers and revenues (total payments received) are monthly total in log after adding one to the level. Any revenues is an indicator for earning positive revenues during the month. Conversion is the conversion rate defined as share of visitors making purchase. Column 3 restrict to observations with sellers earn revenues during the month. Column 5, 7 and 8 restrict to observations where sellers have some visitors during the month. All regressions include cohort, relative month and main sector fixed effect as described in equation 1. Dependent variable means calculated with sellers in the control group. Standard errors clustered by seller. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 3: Treatment Effect on Sellers' Strategies

Variable	Treatment	Dep Var Mean	Variable	Treatment	Dep Var Mean
Market Participation			**Marketing**		
Post Products	0.001	0.34	Paid Ads (Product Counts)	0.002**	0.07
	(0.001)			(0.001)	
Paid Deposits	0.001	0.31	Paid Ads (Traffic Share)	0.001*	0.03
	(0.001)			(0.0004)	
			Promotion	0.0002*	0.0008
Service				(0.0001)	
Active Time (min)	-0.078	19.17			
	(0.128)		**Pricing**		
Reply Time (sec)	-67.725*	23341	Avg. Price Per Buyer	0.005	4.39
	(34.581)			(0.006)	
Conversion Rate	0.0014***	0.1	Avg. Price Per Product	0.0003	3.83
	(0.0005)			(0.001)	

Notes: Table presents estimated coefficients β on treatment assignment dummy with specification 1. Standard errors clustered by seller. All regressions include month, entry date and main industry fixed effect. Post products and paid deposits are indicators for having any active product postings or having put down some deposits during the month. Active time is total number of minutes that sellers' account is active and can answer customer inquiries. Reply time is number of seconds average customers weighted to hear responses from sellers when making inquiries. Conversion rates is measured as share of consumers making purchases among those who made inquiries. Paid ads (product counts) is number of products participating in paid-for-clicks ads. Paid ads (traffic share) is number of consumers visiting sellers' sites from paid channels in log scale (including through paid for clicks ads and other channels). Promotion is number of times sellers participate in the limited time promotional events that the platform regularly organize. Average price per products and average price per buyers are are seller-level prices measured in log scale. Sellers do not have a price measure if they have zero orders. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 4: Interacting New Sellers and Consumers' Purchase

	Dependent variable: Purchase				
		Purchase		Log Spending	Log Order Size
	Same Day	In 3 Days	In a Week	Same Day	
	(1)	(2)	(3)	(4)	(5)
% Treated Seller	0.019***	0.021***	0.024***	0.065**	0.069**
	(0.006)	(0.006)	(0.006)	(0.026)	(0.027)
% Control Seller	0.021***	0.021***	0.027***	0.116***	0.119***
	(0.005)	(0.005)	(0.005)	(0.021)	(0.022)
Incumbent Mean	0.18	0.21	0.22	0.76	0.78
Treatment - Control	-0.0023	0.00011	-0.0027	-0.51*	-0.5
	(0.0066)	(0.0071)	(0.0074)	(0.03)	(0.032)
Observations	1,381,273	1,381,273	1,381,273	1,381,273	1,381,273
R^2	0.680	0.668	0.657	0.698	0.691

Notes: All regressions include search keywords-date-size of consideration set fixed effects, consumer fixed effects and control for average sellers' price level, ratings and number of products offered as well as consumers' baseline characteristics following equation 2. Purchase are dummies for consumers purchasing from some sellers in the consideration set on the day of visit, within 3 days and within a week of visit. Spending is total payments made and order size is the size of order before applying discounts. The later is the main performance metrics for sellers on the platform. For 10% of the cases, consumers placed orders but did not complete the payments. The bottom rows present t-test for listed coefficients with standard errors. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 5: Purchase and Spending within Consideration Sets

	Dependent variable:					
	Purchase				Log Spending	
	Same Day		In a Week		Same Day	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated Seller	0.016***	0.010***	0.010***	0.014***	0.074***	0.050***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.015)	(0.016)
Control Seller	0.009***	0.003	0.007**	0.010***	0.039***	0.010
	(0.003)	(0.003)	(0.003)	(0.003)	(0.012)	(0.013)
Incumbent Mean	0.17	0.17	0.21	0.21	0.71	0.71
Treatment - Control	0.0063	0.0071*	0.0026	0.0041	0.035*	0.039**
	(0.0041)	(0.0041)	(0.0045)	(0.0044)	(0.0019)	(0.0019)
Seller Controls	No	Yes	No	Yes	No	Yes
Observations	300,273	300,273	300,273	300,273	300,273	300,273
R^2	0.119	0.126	0.076	0.084	0.145	0.149

Notes: Sample restricted to set of consumers appearing in at least 2 sets, sets with at least one new sellers and sets where consumers purchased from some sellers in the set within a week. Column 2, 4 and 6 include seller's ratings, price level and number of products listed. Purchase are dummies for consumers purchasing from some sellers in the consideration set on the day of visit or within a week of visit. Spending is total payments made. All regressions include consideration set fixed effects. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 6: Effect on Quality of Purchase

	Dependent variable:		
	Refund Same Order	Return Same Order	Repeat Purchase In a Month
	(1)	(2)	(3)
Treated Seller	-0.005	0.0001	0.020
	(0.018)	(0.012)	(0.016)
Control Seller	-0.027*	-0.015	0.016
	(0.015)	(0.009)	(0.013)
Incumbent Mean	0.072	0.025	0.051
Treatment - Control	0.022	0.015	0.0049
	(0.021)	(0.014)	(0.02)
Control	Yes	Yes	Yes
Observations	54,936	54,936	54,936
R^2	0.950	0.939	0.943

Notes: Sample restricted to set of consumers appearing in at least 2 sets, sets with at least one new sellers and consumer-seller pairs in which case consumers actually placing orders. Control variables include consumer's recent spending, searching and experience as well as seller's ratings, price level and number of products listed. All regressions include consideration set fixed effects. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 7: Estimated Sellers' Type ξ

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
All Sellers	52,241	0.141	0.314	−2.828	0.100	0.100	5.024
By Purchase Status							
No Purchase	37,617	0.098	0.051	−2.828	0.100	0.100	0.101
Has Purchase	14,624	0.252	0.573	−1.965	0.100	0.100	5.024
By Sellers' Type							
Control	3,366	0.136	0.296	−1.445	0.100	0.100	4.939
Treatment	988	0.150	0.353	−0.853	0.100	0.100	3.668
Incumbent	47,887	0.142	0.314	−2.828	0.100	0.100	5.024
By Type Among Sellers with Purchase							
Control	788	0.258	0.593	−1.370	0.100	0.100	4.939
Treatment	252	0.296	0.679	−0.853	0.100	0.100	3.668
Incumbent	13,584	0.251	0.570	−1.965	0.100	0.100	5.024

Notes: Distribution of estimated ξ_j on a sub-sample with 3000 new sellers and the associated incumbents. See appendix B.1 for details. Sellers are grouped based on whether or not at least one consumers have made purchase from these sellers.

Table 8: Sellers' Characteristics and Estimated Type ξ

	Dependent variable:		
	ξ_j		
	Full	Purchase	No Purchase
	(1)	(2)	(3)
Control	0.015	0.142***	0.021***
	(0.010)	(0.032)	(0.003)
Treatment	0.017*	0.109***	0.013**
	(0.019)	(0.055)	(0.005)
Constant	0.096***	0.615***	-0.110***
	(0.003)	(0.007)	(0.001)
Observations	52,241	14,624	37,617

Notes: Distribution of estimated ξ_j on a sub-sample with 3000 new sellers and the associated incumbents. See appendix B.1 for details. Sellers are grouped based on whether or not at least one consumers have made purchase from these sellers.

Table 9: Welfare and Market Share Decomposition

	Revenue	CS	Market Share		Market Expansion		Market Allocation	
			Treatment	Control	Treatment	Control	Treatment	Control
Restrict Access to Training Participants								
Match RF	-0.05	-0.07	-7.71	0.17	-10.87	0.09	-4.94	3.17
All	-0.06	-0.16	-34.68	1.06	-44.66	0.62	-26.05	14.41
Welfare of the Training: Restrict Access to New Sellers								
	-0.19	-0.07	-7.26	-8.76	-8.07	-7.91	0.67	-0.95

Notes: Welfare and market share calculated with a random sample of 60,000 consumers. See B.3 for details. Results presented in the table are percentage difference in comparison to baseline level where traffic assignment is determined by predicted traffic only. Revenues and consumer surpluses are changes from baseline level revenues and consumer surpluses respectively. More market expansion, we calculate the share of revenues coming from consideration sets with some treated new sellers, some control new sellers or only incumbents and calculate market shares of these sets under different matching rules. Results presented here are percentage difference in market share from the baseline. Market allocation is calculated as share of revenues from treated new sellers, control new sellers and incumbents that appear in the consideration sets that include some new sellers. Results presented here are changes in market share comparing to baseline market share.

Figure 1: Screenshot of Training Module and Tasks on Seller's Portal App

(a) Training's Widget (b) Task Details

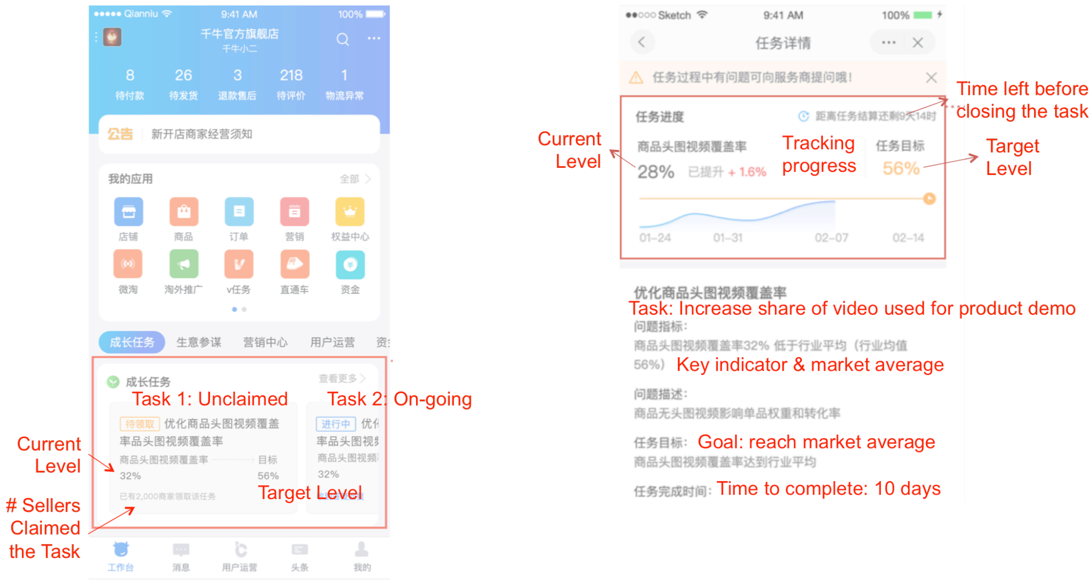


Figure 2: Experimental Sample and Treatment Assignment Over Time

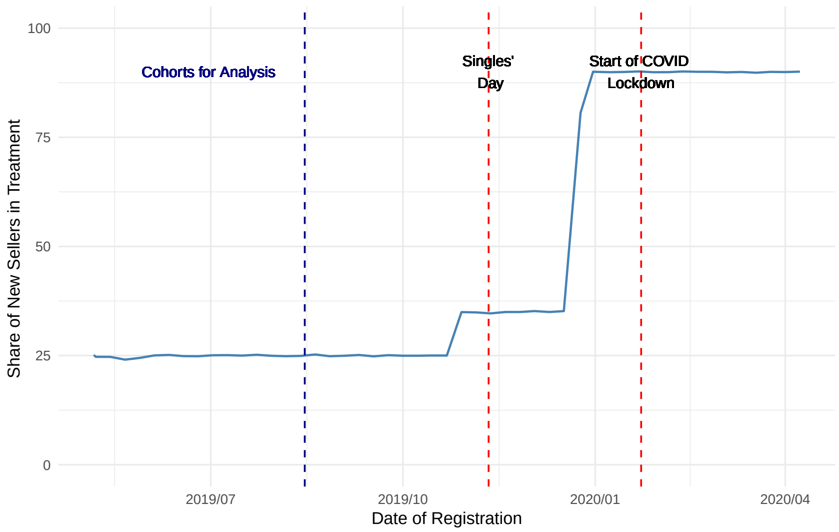


Figure 3: Seller Retention on Training Participation

(a) Unconditional (b) Conditional On Previous Month

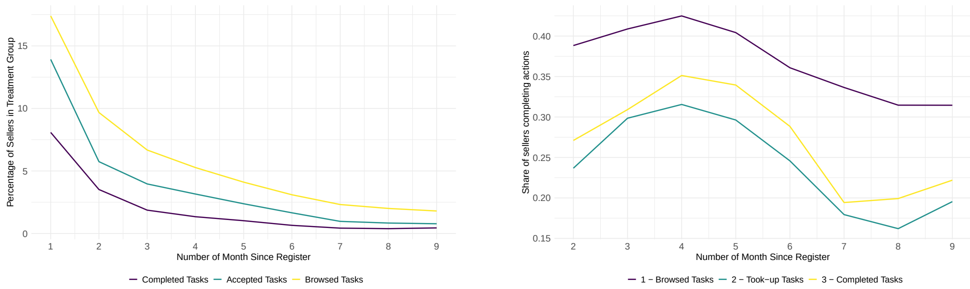
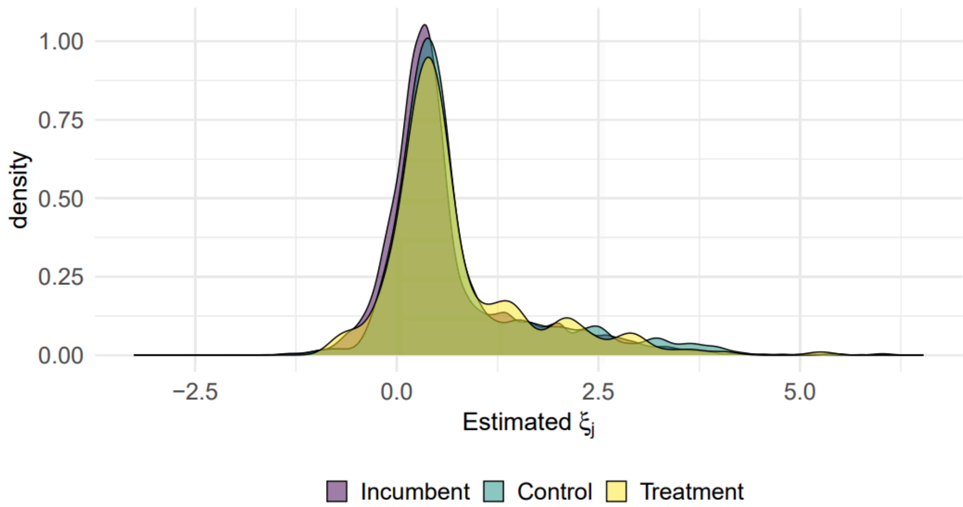


Figure 4: Distribution of Estimated ξ_j



A Additional Empirical Results: Impacts on New Sellers

A.1 Temporal Variations

To explore variations of the treatment effect for sellers at different stages, we estimate the following specification on the balanced sample:

$$Y_{imcs} = \sum_{m=1}^{9} \beta^m Treatment_i M_{im} + \alpha_m + \alpha_c + \alpha_s + \epsilon_{imcs} \tag{8}$$

where M_{im} is the set of indicators for month $m = 1, ..., 9$. We focus on the set of coefficient $\{\beta^m\}_{m=1}^{9}$ that captures the effect of having access to the training during a particular month m since entering the platform. The specification again controls for month of entry α_m, registered cohort α_c and initial sector affiliation α_s fixed effects. Standard errors are clustered at seller level.

The treatment effect on the performance is relatively short-lived. Figure A3 presents the estimated coefficients $\{\beta^m\}_{m=1}^{9}$ in specification (8) on monthly revenues, traffic, conversion rate and number of consumers making purchases. The treatment effect of the training access is positive and significant during the 2nd, 3rd and 4th months on traffic and revenues, but remain insignificant for conversion rate. However, the magnitude of estimated coefficients does not differ significantly from month to month. IV results with actual task take up as the first stage variable follow similar temporal patterns, where the impacts are the strongest during the second and third month on traffic and revenues. The pattern of treatment effect on revenues is consistent with timing of sellers' participation in the training: most of the training participants took up tasks during the initial months (as described in section 2.4) and there could be a lag before actions induced by training become effective

A.2 Heterogeneity

Is the training particularly effective for certain types of new sellers? We examine the heterogeneous impacts of the training by sellers' characteristics in the baseline. Because the content of the training mainly target basic operations and marketing, we expect sellers with limited previous exposure to e-commerce benefit more from the training as the training helps to close the knowledge gap. We characterize sellers from the following dimensions: registration type, gender (if registered as individuals), locations, whether any products were posted on the first day and whether the store is registered as a B2C store on the first day[44]. Since no sellers took up tasks on the first day of entry, we consider listing products and sellers' registration type on the first day as part of pre-treatment characteristics. We then estimate the following specification

$$Y_{imcs} = \beta Treatment_i + \gamma Type_i + \delta Treat_i \times Type_i + \alpha_m + \alpha_c + \alpha_s + \epsilon_{imcs} \qquad (9)$$

where as before $Treat_i$ denotes assignment to treatment or control group and $Type_i$ specify whether sellers have characteristics aforementioned. δ captures the heterogeneous treatment effects on different types of sellers. Table A6 summarizes the results on log revenues by sellers' types. Overall, estimated β in these specifications have similar magnitudes as estimates using equation 1, but there are no differential treatment effects by sellers' types, gender, actions on the first day of entry on revenues. Slightly surprising result is that there is no differential treatment effect for sellers with different level of preparedness. Comparing to the rest, sellers who post products on the very first day of entry could be better prepared or are more experienced. Hence, these sellers might find the basic part of the training less useful, yet we do not find such results[45].

To evaluate the impacts of offline business environment, we group the sellers based on their registered locations[46]. Table A7 presents the results on heterogeneous treatment effect by sellers' locations on log monthly revenues. γ captures average performance of sellers in different part of the country comparing to those coming from the remaining parts. There are significant variations in average performance for sellers from different parts of the country. Sellers from the coastal southern provinces significantly out-perform the rest while those coming from the less-developed western part of the country lagged behind. Performance of sellers in different regions is consistent with economic development in the offline world. Training is less helpful for sellers from less developed regions, as these sellers are less likely to take-up the training. Therefore, even the training program offers the exact same materials to all sellers, sellers coming from less pro-business areas are less likely to take advantage of such knowledge. As a result, the training does not help those lagging behind to catch up, but instead further strengthens the competitive edge of new sellers from more developed regions.

[44]To register as a B2C store, potential sellers must obtain formal approval from the platform. The minimum requirements include having a brand name and a formally registered firm. 97% of sellers in the sample are registered as C2C stores. C2C stores can be converted to B2C stores later on. Among all sellers that eventually become B2C stores, 66.7% of them converted later on and sellers in the treatment group are more like to convert.

[45]We also do not find training to be more useful for sellers who post products after the first day of entry.

[46]The location information on ID cards for individual sellers may not reflect where the sellers actually reside at the moment because the location indicates ID card holder's birthplace, rather than current residence. The internal migration patterns implies that we are under-counting sellers living in the coastal provinces as these provinces are major destinations of migration. Similarly, firm's registered locations might not be the same as where the firms actually operate in, but in this case the direction of the bias is unclear.

B Details on Structural Estimation

B.1 Data and Sampling

The sample we use for the structural estimation of consumer demand is an adapted version of the sample used in 4.2. The population is the consumer-seller pair sample where each consumer i seller j pair belongs to a search query - efforts - date combination s. The difference from the sample used in 4.2 is that we also include consideration sets that contains incumbents only, as opposed to just new sellers. Because of the computational constraint and setup of the model, we are unable to estimate the model on the full sample. Instead, we use the following approach to construct the sample:

1. Randomly sample 3,000 new sellers (treated and incumbents) from the pool of new sellers that appear in the full consumer-seller pair sample.
2. For each new sellers, obtain all the incumbents sellers that appear in the same consideration sets s as these new sellers do, include these sellers to the estimation sample.
3. For all the incumbents who appear in the same sets with the new sellers, obtain all the consideration sets these sellers appear in as well as other incumbents that appear in the same sets as they do, add these sellers to the estimation sample.
4. Iterate the previous step until all the new sellers and incumbents in the estimation sample appear in at least two consideration sets.

We require all the sellers in the estimation sample to appear in at least two consideration sets because otherwise the ξ_j would not be identified. As a result, our final sample consists of sellers who have higher traffic shares because these sellers are more likely to appear in multiple consideration sets. We use traffic to refer to number of visitors a seller obtain within a 30-day period. The final sample consists of 52,241 sellers and 1,312,967 observations in 323,584 consideration sets. 3,366 sellers are in the control group and 988 sellers belong to the treatment group. We have more new sellers than what we originally sampled because additional new sellers are incorporated into the estimation sample in the iteration process. 18.1% of the observations are selected as part of the estimation sample.

The estimation sample is at seller level, even though in the search process, consumers access specific sellers' sites by searching specific products rather than front page of the sites. Sellers almost always offer multiple products and sometimes could span different sectors. Consumers could purchase multiple products from the sellers they visited and in particular they could purchase products other than they ones that direct them to sellers' sites. Since we do not observe consumers' browsing history in the stores, we are unable to fully capture such process. Instead, since we are predominantly interested in the seller level characteristics, we aggregate all the purchase and browsing behaviors to consumer-seller level, rather than to consumer-seller-product level. Number of products offered and ratings are measured at seller level. Pricing is the quantity weighted prices of all products that the sellers offer.

The instruments we used for sellers' pricing and number of products offered are variables capture intensity of the platform's rule enforcement. The variables are constructed as averaging over all the sellers belonging to the same sector during the past 30 days period weighted by each seller's traffic. We subtract seller j's own weighted measure from the weighted mean. These instruments include frequencies of identified and enforced rule violations and share of sellers identified as selling fake or counterfeit products or as boosting sales with fake orders. We standardized these variables to have mean zero and unit standard deviations for estimation.

B.2 Estimation Details

We modify the baseline model as described in 5.2. In the actual estimate, we use **??** to approximate the probability of consumer i choosing a specific seller j in the consideration set, which gives the probability parameter in a Bernoulli distribution. We use the simulated maximum likelihood to identify the parameters of interests by matching with realized purchases on the day of visits. In the baseline model, we include the sector specific fixed effect ξ_s but keep the price coefficient α constant for all sectors. We enrich the model with sector specific coefficient α_s and the estimated results are similar. To account for endogeneity of product offerings and pricing, we use instruments mentioned above. We model price level and log number of products offered following multinormal distribution where the respective means are determined by equation 5. We jointly estimate the first stage for strategies with the consideration set based demand.

To estimate the matching rule, we use the new seller sample described in **??** where we re-define month relative to the time when sellers first post the products. As described in 5.2, we include dummies for having no visitors or zero conversion rate in the previous periods. The traffic measure is converted to log scale. We use a flexible polynomials for $f(\mathcal{T}_{jt-1}, C_{jt-1}, C_{jt-2})$ and test for the changes in R^2 when adding higher order terms of $\mathcal{T}_{jt-1}$, C_{jt-1}. We also add relative month, calendar date and initial sector fixed effects. Adding higher order terms of lagged traffic and conversion does not significantly improves precision of predicted traffic. Excluding the fixed effects will reduce R^2 by construction but the impacts on predicted traffic is small. The most important predictor is the lagged traffic and the relationship between current period traffic and previous period traffic is close to linear. To test the precision of the prediction, we use cross validation method and calculate the average residuals on the training sample. Table A11 shows measures of prediction precision on current period traffic with different specifications. For the actual estimate to generate predicted traffic, we use estimated coefficients on $\mathcal{T}_{jt-1}$ and C_{jt-1} and C_{jt-2} with linear specification without fixed effects.

B.3 Counterfactual Details

To run the counterfactual analysis, we randomly sample 60,000 consumer - search keyword - date combinations and obtain their corresponding number of sellers visited. We construct the potential pool of sellers that a specific consumer searching a particular keyword could sample from as **all** the sellers who were visited by any consumers searching that keywords on that date. For each seller in the pool, we calculate their predicted traffic $\hat{T}_j$ using the estimated $f(\cdot)$ as described in the previous section using lagged traffic and lagged conversion rates. The predicted traffic is a good approximation of the actual traffic these sellers acquire. Ideally, we should use query-specific predicted traffic as the sampling weights, but such data is not currently available. The sampling weights for seller j in a specific query-date pool is given by $w^0_{js} = \frac{\hat{T}_j}{\sum_{k \in S} \hat{T}_k}$. We use w^0_{js} are the baseline sampling weights. With the estimated sampling weights for all the sellers in the pool, we randomly sample sellers from consumer-set specific pool where number of sellers is the same as number of sellers that the consumer actually visited during in that particular search sessions. Therefore, the only part that is changed in the counterfactual analysis is the composition of sellers in consumers' consideration sets, while size of the sets and which keywords consumers searched are all kept constant.

For the counterfactual analyses, we adjust sellers' sampling weights to consider the welfare impacts of the current training and the impacts of scaling up the training to cover all the new sellers. To quantify the impacts of the training, we make the assumption that treated new sellers should have the same behaviors as the control new sellers. Among all the sellers that appear in the sampling pooling, 5.9% of them are control new sellers and these sellers account for 3.2% of observations in the sampling pool. Currently, 1.74% of sellers in the pool are treated new sellers and they make up for 1.07% of the appearance. Without the training, treated new sellers should make up for similar share of the observations in the sampling pool as the control new sellers do, in which case they should make up for 0.94% of the observations, which is a 12.2% drop from their current shares. If we further assume that only training participants are subject to the influence of the training, and since that training participants make up for 51.16% of the treated new sellers, their appearance in the sampling pool should be dropped by 24.66%. Therefore, in the counterfactual analysis, we randomly drop 24.66% of training participants' appearance from the sampling pool and recalculate the sampling weights of other sellers in the pool after removing consumers' access to these sellers. With the new sampling pool and the updated weights, we reconstruct consumers' consideration sets and calculate the impacts on consumer surplus and sellers' revenues following the specifications in section 5.4. To decompose the changes in sellers' revenues, we compute the market share of three types of consideration sets as well as different types of sellers' market shares if they appear in the consideration sets with some new sellers under each counterfactual scenarios. We then calculate the changes in market shares from the baseline market allocation.

To quantify the welfare impacts of scaling up the training to cover all the new sellers, we adjust sellers' sampling weights by lowering some new sellers' probabilities to appear in consumers' consideration sets by changing the their sampling weights to a small ϵ. We randomly sample 9.7% of new sellers among all of those appearing in the potential pool and change their sampling weights in the corresponding sampling pool to ϵ to match with what we find in the reduced-form effect on traffic. Weights of other sellers that appear in the same pools as these sellers do are re-calculated accordingly. We construct the alternative consideration sets based on these new sampling weights and calculate the consumer surpluses as well as sellers' revenues.

C Online Sellers' Survey

We conducted an online survey in August 2019 with sellers to gather some basic demographics information and their opinion about the training. The sampling was stratified by sellers' engagement with the training and we over-sampled sellers who were more involved with the training, i.e. sellers who took up more tasks. In the end with collected 566 responses. Detailed results are presented in table A13. Since most of the respondents are training participants, they may not form a representative sample of sellers on the platform. These respondents are likely to be more active and have higher sales. Moreover, compared to typical sellers, these sellers appear to have higher than average ownership of manufacturing factories (32.4%) and offline stores (18.9%).

The survey shows that even among the training participants, sellers differ in terms of their background, experience, education and financial resources. However, while vast majority of the active new sellers are small and inexperienced, a substantial share of them are reasonably educated and express clear interests to participate in the e-commerce. Results from the online survey show that 71.9% have 1 or 2 employees, 74.3% have no or less than one year of experience in e-commerce, but 67.3% have completed at least high school education. About 58.8% of sellers in the sample report that they intend to make running the e-commerce store as their main job and 48.2% have invested more than 10,000 RMB ($1430) into their online businesses. The platform does not have a systematical approach to collect demographic data of from the sellers other than those collected during the registration[47].

[47] We could potentially gather more information such as predicted education, income level and total spending on the platform through the affiliated financial subsidiaries.

D Additional Tables and Figures

Table A1: Treatment Effects on Monthly Revenues

	Dependent variable:					
	Monthly Revenues					
	Indicator	Raw	Log	Winsorized		
				99th	99.5th	99.9th
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.002**	-939.245	0.017***	25.436**	26.175	85.084
	(0.001)	(1,447.148)	(0.006)	(12.478)	(22.308)	(54.321)
Dep Var Mean	0.19	7018.16	1.39	1322.79	2019.61	3472.93
Observations	6,409,062	6,409,062	6,409,062	6,409,062	6,409,062	6,409,062
R^2	0.152	0.0001	0.132	0.027	0.024	0.019
Adjusted R^2	0.152	0.00004	0.132	0.027	0.024	0.019

Notes: Dependent variables are total revenues in the seller sample. All regressions include cohort, initial sector and relative month fixed effect. Dependent variable means calculated with sellers in the control group.* significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table A2: Treatment and Speed of Listing Products

	Dependent variable:				
	Number of Days Passed Before Listing First Items				
	OLS	OLS	IV	IV	OLS
	(1)	(2)	(3)	(4)	(5)
Treatment	0.057	0.067			
	(0.144)	(0.207)			
Took-up Tasks (Fitted)			0.835	0.664	
			(2.121)	(2.046)	
Took-up Tasks					7.180^{***}
					(0.597)
Sample	All	Late	All	Late	Late, Treatment
Dep Var Mean	16.05	23.99	16.05	23.99	24.03
Observations	476,292	318,792	476,292	318,792	79,477
R^2	0.001	0.002	0.001	0.002	0.004
Adjusted R^2	0.001	0.001	0.001	0.002	0.003

Notes: Sample restricted to sellers who have posted at least one product during the sample period. Dependent variable is number of days passed since registration before sellers posting the first product. Column 2, 4, and 5 further restricted sample to sellers who posted products on the second day or later. Column 5 again restricts the sample to sellers who have listed products on the second day or later and are assigned to the treatment group. For instrumental regressions in column 3 and 4, the instrument is being assigned to the treatment group. All regressions include cohort fixed effect. Dependent variable means calculated with sellers in the control group. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table A3: IV Results on Main Outcomes

	Dependent variable:					
	Take-up Tasks	Any Revenues	Revenues	# Visitors	# Buyers	Conversion Rate
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.257***					
	(0.001)					
Take-up Any Tasks		0.006**	0.066***	0.052**	0.032***	0.0001
		(0.003)	(0.024)	(0.021)	(0.012)	(0.001)
Dep Var Mean	0	0.19	1.39	1.73	0.57	0.04
Observations	6,409,062	6,409,062	6,409,062	6,409,062	6,409,062	2,593,762
R^2	0.232	0.153	0.133	0.208	0.107	0.049
Adjusted R^2	0.232	0.153	0.133	0.208	0.106	0.048

Notes: The first stage variable an indicator for whether or not sellers have taken up at least one task during the nine-month period. Column 1 presents the estimated on first stage outcome with treatment assignment the instrumental variable. All specification are 2-stage least square results using treatment assignment as the instrument.Traffic (number of visitors), number of buyers and revenues (total payments received) are monthly total in log after adding one to the level. All regressions include cohort, relative month and initial industry fixed effect. Standard errors clustered at seller level. Dependent variable means calculated with sellers in the control group. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table A4: OLS Results on Main Outcomes

	Dependent variable:							
	Log Revenues	Any Revenues	Revenues	Log # Visitors		Log # Buyers		Conversion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Take-up Tasks	1.847***	0.223**	0.771***	1.735**	0.703***	0.825***	0.686***	0.005***
	(0.017)	(0.002)	(0.019)	(0.014)	(0.015)	(0.009)	(0.012)	(0.0004)
Dep Var Mean	1.39	0.19	7.14	1.73	4.29	0.57	1.4	0.04
Sample	Full	Full	Earn Revenues	Full	Have Visitors	Full	Have Visitors	Have Visitors
Observations	1,593,234	1,593,234	314,376	1,593,234	646,894	1,593,234	646,894	646,894
R^2	0.193	0.204	0.108	0.274	0.128	0.160	0.098	0.048
Adjusted R^2	0.193	0.204	0.1087	0.274	0.128	0.159	0.097	0.047

Notes: Sample restricted to sellers with access to training. Main explanatory variable is having take-up at least one tasks during the sample period (not taking up tasks during the month). Traffic (number of visitors), number of buyers and revenues (total payments received) are monthly total in log after adding one to the level. All regressions include cohort, relative month and initial industry fixed effect. Standard errors clustered at seller level. Dependent variable means calculated with sellers in the control group. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table A5: Treatment Effect on Sellers' Ratings and Refunds

Variable	Treatment	Dep Var Mean	Variable	Treatment	Dep Var Mean
Ratings			**Refunds and Reviews**		
Products	0.004	1.25	% Refund (Amount)	-0.0004	0.21
	(0.004)			(0.001)	
Service	0.004	1.26	% Complaints	-0.003	0.05
	(0.004)			(0.006)	
Logistics	0.005	1.26	Rule Violations	0.0002	0.22
	(0.004)			(0.001)	
			% Good Reviews	-0.00001	0.99
				(0.0002)	

Notes: Table presents estimated coefficients β on treatment assignment dummy with specification 1. Standard errors clustered by seller. All regressions include month, entry date and main industry fixed effect. Ratings are customer ratings variables that the platform calculate and assigned to sellers based on customers' reviews and ratings. The ratings scale between 0 to 5, on the dimensions of accuracy of product descriptions, quality of customer service and logistics. % refunds calculated as total refunds requested over total payments made. % complaints defined as number of complaints over total number of orders. Rule violations is frequency that sellers violate the platform's rules, see more details on that rule violations mean in B.1 the description of instruments. % of good reviews are share of good reviews out of all reviews. Vast majority of the reviews are positive. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table A6: Treatment Effect Heterogeneity on Sellers' Basic Types

	Dependent variable: Log Revenues				
	Registration Type		Post Products		B2C Sellers
	Female	Firm	First Day	Later Days	
	(1)	(2)	(3)	(4)	(5)
Treatment	0.016*	0.009	0.014**	0.011*	0.016**
	(0.008)	(0.006)	(0.007)	(0.006)	(0.006)
Seller Type	-0.494***	1.504***	-0.130***	1.265***	3.842***
	(0.006)	(0.009)	(0.008)	(0.008)	(0.050)
Treatment × Seller Type	0.003	0.015	0.013	0.014	0.077
	(0.012)	(0.017)	(0.015)	(0.013)	(0.098)
Dep Var Mean	1.39	1.39	1.39	1.39	1.39
Observations	6,409,062	6,409,062	6,409,062	6,409,062	6,409,062
R^2	0.138	0.174	0.132	0.163	0.147
Adjusted R^2	0.138	0.174	0.132	0.163	0.147

Notes: Standard errors clustered by seller. All regressions include month, cohort and initial sector fixed effect. Dependent variable is monthly revenues in log scale after adding one to base level. The interaction variables are indicators for whether or not sellers are females, are registered as firms, post products on the very first day of entry or during some later days and lastly whether or not sellers register as B2C sellers. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table A7: Treatment Effect Heterogeneity by Sellers' Registered Location

	Dependent variable: Log Revenues					
	Beijing Vicinity	Resource-Oriented	Northeast	Coastal South	Central	West
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.015**	0.018***	0.019***	0.014*	0.015**	0.019***
	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Location	-0.039***	-0.371***	-0.273***	0.470***	-0.128***	-0.642***
	(0.008)	(0.013)	(0.012)	(0.006)	(0.007)	(0.007)
Treatment × Location	0.011	-0.042	-0.048*	0.006	0.008	-0.013
	(0.017)	(0.026)	(0.025)	(0.013)	(0.014)	(0.014)
Dep Var Mean	1.39	1.39	1.39	1.39	1.39	1.39
Observations	6,409,062	6,409,062	6,409,062	6,409,062	6,409,062	6,409,062
R^2	0.132	0.132	0.132	0.138	0.132	0.136
Adjusted R^2	0.132	0.132	0.132	0.138	0.132	0.136

Notes: Standard errors clustered by seller. All regressions include month, entry date and main industry fixed effect. Dependent variable is monthly revenues in log scale. Indicators are sellers registration locations clustered into different regions. Beijing Vicinity includes Beijing, Tianjin, Heibei and Shandong; resource-oriented provinces include Shanxi, Neimenggu, Gansu and Ningxia; northeastern provinces are Heilongjiang, Jilin and Liaoning; coastal southern provinces are Jiangsu, Shanghai, Zhengjiang, Fujian, Guangdong and Hainan; central provinces are Anhui, Jiangxi, Henan, Hubei and Hunan; western provinces are Tibet, Xinjiang, Yunnan, Guangxi, Sichuan, Chongqing, Guizhou, Shaanxi and Qinghai. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table A8: Summary Statistics: Consumer-Search Session Sample

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Number of Sellers Browsed	1,381,273	4.892	2.691	3	3	6	55
Share of Treated New Sellers	1,381,273	0.016	0.065	0	0	0	1
Share of Control New Sellers	1,381,273	0.023	0.082	0	0	0	1
Purchase (Same Day)	1,381,273	0.188	0.390	0	0	0	1
Purchase (in 3 Days)	1,381,273	0.215	0.411	0	0	0	1
Purchase (in A Week)	1,381,273	0.231	0.422	0	0	0	1
Pay Amount	1,381,273	34.282	151.451	0	0	0	14,649
Order Size	1,381,273	40.865	227.007	0	0	0	88,000
Recent Spending	1,381,273	1,822.888	6,695.639	0	9.9	713	1,960,667
Recent Search	1,381,273	184.069	4,370.249	0	31	165	1,580,514
Consumers' Experience	1,381,273	4.983	2.378	−2	4	7	14
Sellers' Price Level	1,381,273	245.912	616.886	0.01	62.861	235.551	69,100
Number of Listed Products	1,381,273	1,196.689	11,734.010	0	100	831	4,216,488
Seller's Rating	1,381,273	13.032	2.916	−1	11.5	15	20

Notes: Table presents the summary statistics of main variables in the consumer-search session sample. Each observation is a consumer-search session.

Table A9: Summary Statistics: Consumer-Seller Sample

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Number of Sellers Browsed	300,273	4.685	2.621	1	3	5	55
Purchase (Same Day)	300,273	0.183	0.387	0	0	0	1
Purchase (in a Week)	300,273	0.219	0.414	0	0	0	1
Pay Amount	300,273	34.29	18.74	0	0	0	10,000
Refund	54,910	0.068	0.252	0	0	0	1
Return	54,910	0.023	0.150	0	0	0	1
Repeat Purchase	54,910	0.049	0.217	0	0	0	1
Recent Spending	300,273	4,886.792	9,656.899	0	141.8	5,691.6	260,495.5
Recent Search	300,273	160.37	194.731	0	48	211	5,133
Price Level (Seller)	300,273	202.96	392.37	0	53.2	221.5	80,000
Number of Listed Products (Seller)	300,273	527.033	7,681.935	0	13	299	3,359,564
Seller's Rating	300,273	11.209	4.059	−2	5	7	12

Notes: Table presents the summary statistics of main variables in the consumer-seller sample. Each observation is a consumer-seller pair where consumers purchased from some sellers during the search session within a week of visit.

Table A10: Visiting New Sellers and Consumers' Purchase

	Dependent variable: Purchase				
		Purchase		Log Spending	Log Order Size
	Same Day	In 3 Days	In a Week	Same Day	
	(1)	(2)	(3)	(4)	(5)
Treated Sellers	0.003**	0.004**	0.004**	0.011*	0.013*
	(0.001)	(0.002)	(0.002)	(0.007)	(0.007)
Control Sellers	0.004***	0.004***	0.005***	0.023***	0.023***
	(0.001)	(0.001)	(0.001)	(0.005)	(0.006)
Incumbent Mean	0.18	0.21	0.22	0.76	0.78
Treatment - Control	-0.00035	0.000088	-0.0005	-0.012	-0.011
	(0.0017)	(0.0018)	(0.0019)	(0.0076)	(0.0081)
Observations	1,381,273	1,381,273	1,381,273	1,381,273	1,381,273
R^2	0.680	0.668	0.657	0.698	0.691

Notes: All regressions include search query-date-size of consideration set fixed effects, consumer fixed effects and control for average sellers' price level, ratings and number of products offered as well as consumers' baseline characteristics. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table A11: Prediction Precision: Traffic

Lagged Traffic Degree	Lagged Conv. Rate Deg	FE	R^2	RMSE	MAE
1	1	Y	0.63	1.63	1.18
1	1	N	0.6	1.7	1.23
2	1	N	0.6	1.69	1.22
2	1	N	0.61	1.68	1.21
3	2	N	0.6	1.69	1.22

Notes: Table shows measure of prediction's precision with different specifications on current period traffic. Precision calculated with on test data.

Table A12: Example of Tasks and Their Classifications

Task	Indicator	Area of Focus	Function	Type
Acquire customers' reviews	reviews	ratings	knowledge	outcome
Acquire free traffics	visitors from search channel	marketing	knowledge	outcome
Choose proper promotion products	payment received	basic	knowledge	outcome
Complete an order	payment received	basic	knowledge	outcome
Expand base of followers	followers	customers	knowledge	outcome
Improve "add to shopping cart"	add to cart	marketing	knowledge	outcome
Improve buyer review section	reviews	basic	reminder	outcome
Improve conversion rate: inqury	conversion	service	knowledge	outcome
Improve conversion rate: make payment	conversion	marketing	knowledge	outcome
Improve fans' engagement	followers	customers	reminder	outcome
Improve payments from returning customers	payments received	customers	knowledge	outcome
Engage with customers via weitao	followers' activities	customers	reminder	action
Improve per consumer spending	avg. order size	marketing	knowledge	outcome
Improve ratings on customer service	ratings	ratings	knowledge	outcome
Improve ratings on product quality	ratings	ratings	knowledge	outcome
Decorate store frontpage on app	decoration	basic	reminder	action
Improve tag/bookmark rates	bookmarked	marketing	knowledge	outcome
Imrpove ratings on delivery	ratings	service	knowledge	outcome
Optimize products' titles	traffics	marketing	knowledge	outcome
Participate in official sales events	sign-up	marketing	reminder	action
Pay security deposits	deposites	basic	reminder	action
Post products on store page	number of products	basic	reminder	action
Setup bonus after purchase	bonus	basic	reminder	action
Setup free return and refund	return policy	basic	reminder	action
Setup free trial / offer free samples	free trial	basic	reminder	action
Setup paid "wangpu"	wangpu	basic	reminder	action
Setup store coupons and discount	coupons	basic	reminder	action
Shorten average time to delivery	delivery time	service	knowledge	outcome
Shorten response time to customer inquries	response time	service	knowledge	outcome
Upload videos for product descriptions	vidoes	basic	reminder	action

Notes: Listed tasks are a subset of all tasks offered to the sellers. Over time service providers also created more tasks and the platform invested in streamlining and regularizing the tasks offered. Tasks are order in sequence of priorities. Each task is triggered by a particular indicator. For the outcome based tasks, comparison are made with other sellers in the same industry. Tasks are classified based on main area of focus, the functions they served and how they are evaluated.

Table A13: Summary of Sellers' Survey

Category	Fraction	Category	Fraction
Respondent Chars		**Business Chars**	
Education		*Sources of Supply*	
Primary	2.8%	Own factory	32.5%
Middle School	29.0%	Offline wholesale markets	19.2%
High School	23.7%	Online whotesale markets	21.7%
Some College	28.0%	Distribution/brand subsidary	19.5%
Bachelors	15.4%	Others	7.1%
Master's and Above	0.8%		
Professional Degrees (e.g. MBA)	0.3%	*Number of Employees (inc. owners)*	
		1 - 2 persons	71.9%
Exp in Retail		3 - 5 persons	21.8%
None	36.7%	6 - 10 persons	3.9%
Less than a year	25.6%	>10 persons	2.4%
1 to 3 years	17.2%		
More than 3 years	20.5%	*Total investments*	
		<5k RMB	32.3%
Exp in E-commerce		5k - 10k RMB	19.5%
None	36.3%	10k - 50k RMB	25.2%
Less than a year	38.0%	50k - 100k RMB	9.3%
1 to 3 years	16.5%	100k - 200k RMB	5.0%
More than 3 years	9.2%	>200k RMB	8.7%
Goal			
No specific goal	3.1%		
As part-time job	19.2%		
As main job	58.8%		
Expand offline business online	18.9%		

Notes: Online survey implemented with users assigned to treatment group for the training intervention in August 2018. Separate messages were sent out based on sellers' engagement with the training defined by number of tasks accepted and whether or not sellers have browsed contents of the training. Survey response rates are higher among sellers that were more engaged in the training. All fractions shown adjusted for the sampling and response rate differential.

Figure A1: Timing of Taking-up First Tasks by Task Contents

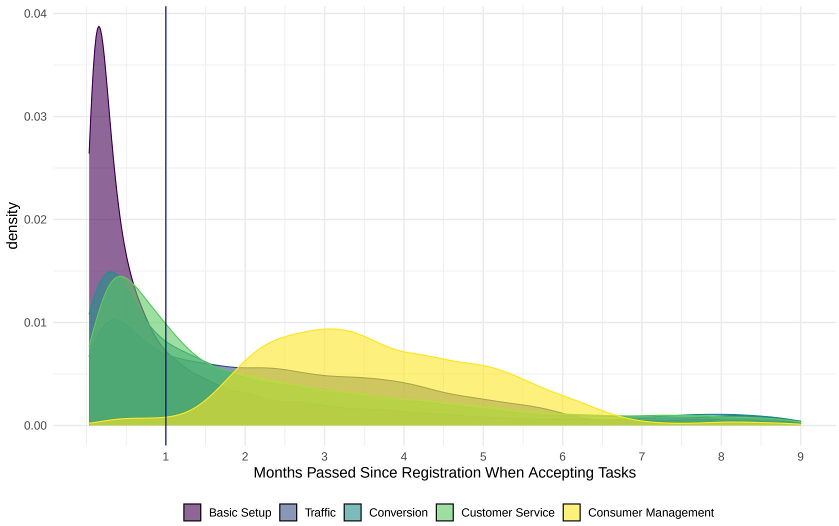


Figure A2: Quantile Treatment Effect on Revenue Over Time

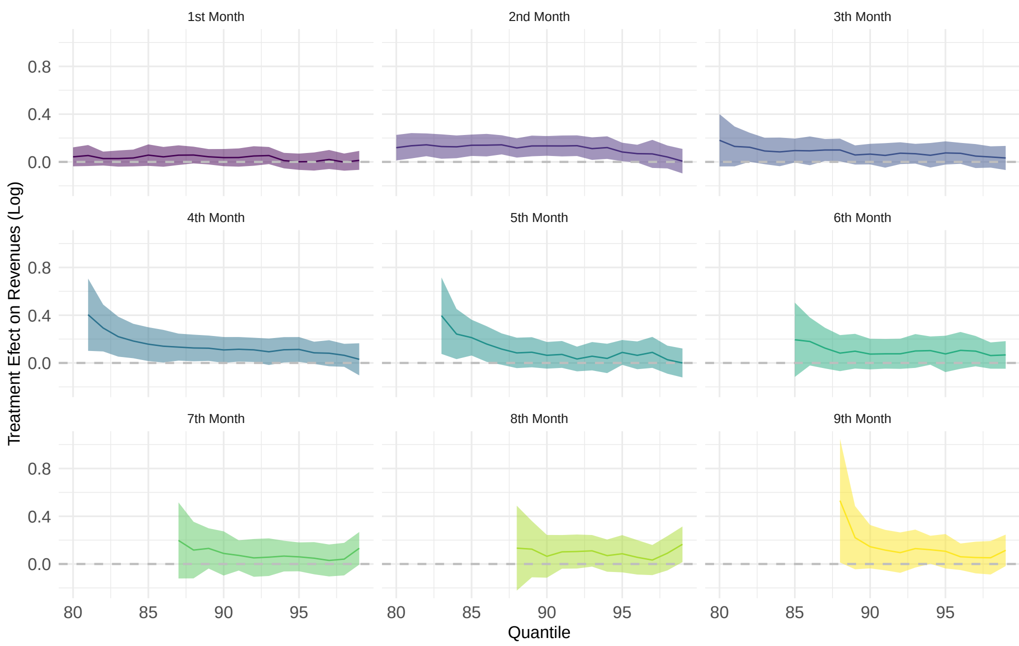
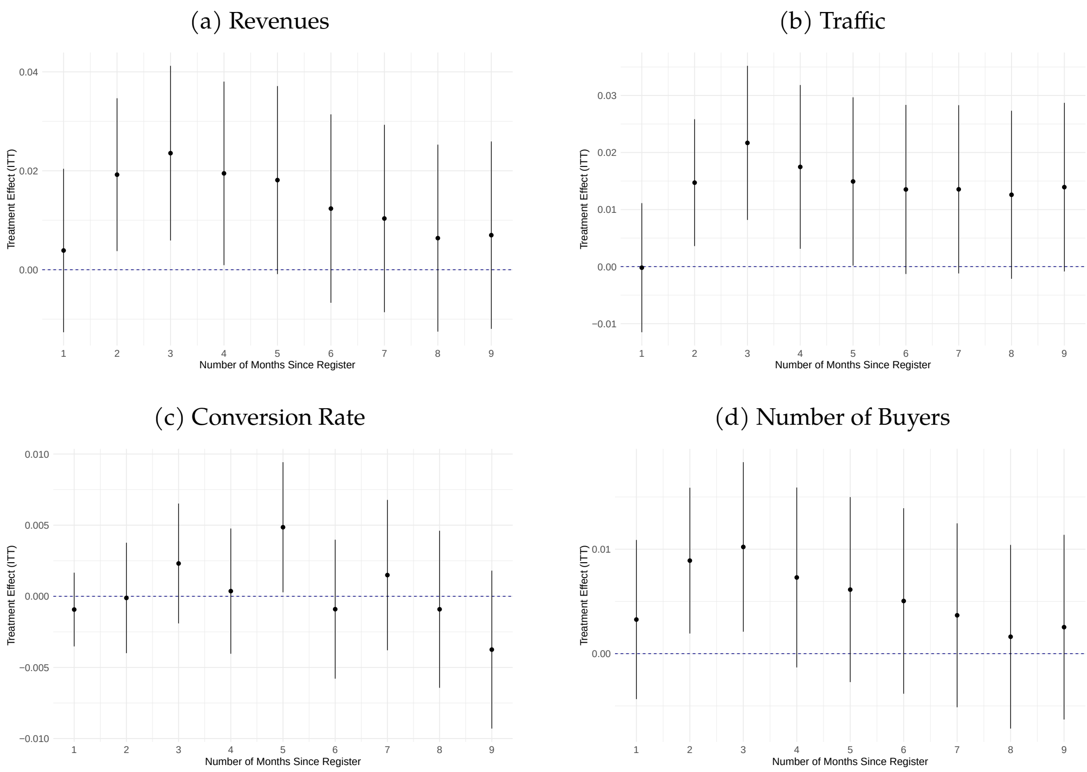


Figure A3: Long Term Treatment Effect on Main Outcomes



(a) Revenues (b) Traffic (c) Conversion Rate (d) Number of Buyers

Figure A4: Distribution of Size of Search Sets

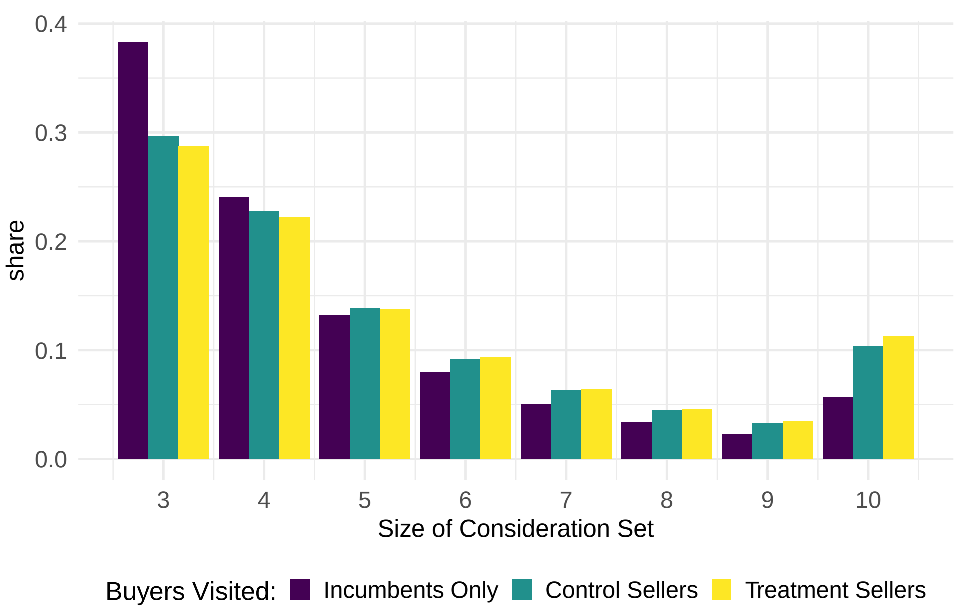
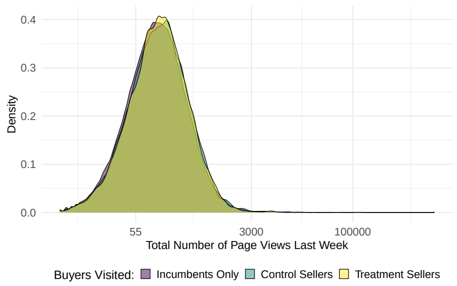


Figure A5: Distribution of Previous Week's Spending and Search Intensity

(a) Total Page Views



(b) Total Spending

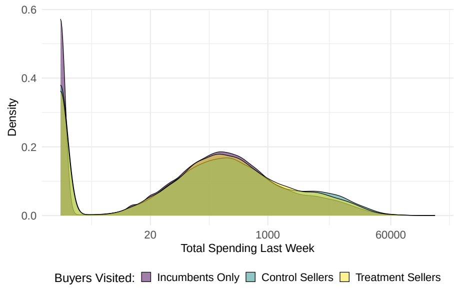
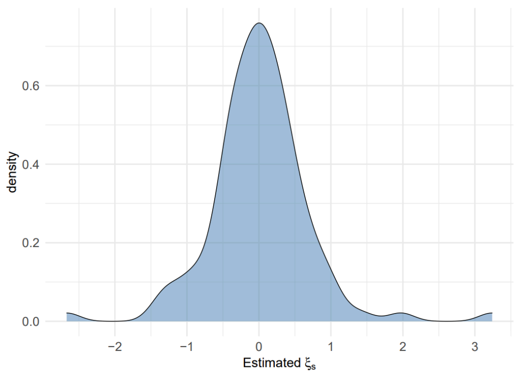


Figure A6: Distribution of Estimated Parameters

(a) ξ_s



(b) Elasticity