The Online Growth Barrier and AI Entrepreneur Training

An Experiment with Two Million Sellers and A Hundred Million Buyers*

Yizhou Jin
UC Berkeley

Zhengyun Sun
Harvard

April 14, 2021

Abstract

We collaborate with a large e-commerce platform, and construct a new dataset of consideration sets by matching consumer search and browsing behavior with seller product and operations information. We also conduct a year-long experiment among new sellers on the platform. The treatment group receives access to a free online entrepreneur training program customized by an AI algorithm. We first document a growth barrier ("GB") in the competition for consumer attention: new sellers must achieve 10% higher conversion rates to obtain the same consumer exposure as incumbents. The training program lowers this barrier by enhancing sellers' marketing and customer service efforts, hence boosting their revenue. We then estimate an empirical model to capture heterogeneity in consumer demand. This quantifies the effect of the GB: conditional on making a sale, new sellers have 23% higher quality than incumbents. By lowering this barrier, the training program benefits buyers. Although only 0.25% of all sellers on the platform are trained, removing the program would have reduced consumer surplus by 0.07%.

Keywords: AI, entrepreneurship, e-commerce, platform, steering, search, consideration set

*Jin: jyz@berkeley.edu; Sun: zsun@g.harvard.edu. We thank Susan Athey, Jie Bai, Emily Breza, Long Chen, Shawn Cole, Asim Khwaja, Gabriel Kreindler, Michael Kremer, David McKenzie, Song Ma, Ariel Pakes, Gautam Rao, Yulu Tang, Yuan Tian, Christopher Woodruff, and various seminar participants. We also thank the product development team at the unnamed e-commerce platform, and our data contacts and providers for making this project possible. Ability to publish is not contingent on results. Please email carolina_harvey@harvard.edu for questions about the data usage agreements.The trial were registered on the AEA Social Science Registry (AEARCTR-0006725). Human Subjects Approval was obtained from Harvard University (IRB19-0647).

1 Introduction

Digital platforms can provide vast market access at low entry cost for new businesses. This has stimulated a large amount of entrepreneurial activities: in 2020, 1.3 million new sellers entered Amazon marketplace,[1] while Youtube hosted seven million new channels.[2] However, concentration on platforms has increased dramatically (Bärtl 2018). Among practitioners, there is growing recognition for a "growth barrier": new businesses have a disadvantage in reaching consumers and thus cannot effectively compete with incumbents. Nonetheless, direct measurement and evaluation of such a barrier is limited.

Conceptually, platforms can directly promote new-business success and reduce the growth barrier by steering consumer attention. However, effective steering requires strong quality signals that are hard to obtain among new businesses (Hagiu and Wright 2020). Bai, Chen, Liu, and Xu (2020) show that small-and-medium exporters online face sizable search friction. Random demand shocks can thus strongly impact growth trajectory. This is similar to Fradkin (2015), Reshef (2019), and Cullen and Farronato (2020), which show that digital platforms may not enjoy increasing returns to scale in terms of matching efficiency due to search friction, particularly between consumers and new suppliers. Tadelis (2016), Lewis and Zervas (2019), and Farronato, Fradkin, Larsen, and Brynjolfsson (2020) show that past consumer choice and reviews may generate large information externalities and form an entry barrier that is hard to capture by other information available to the platforms. Sahni and Nair (2020) and Moshary (2021) show that selection into paid advertising can signal seller quality, but few new businesses advertise even at zero cost (Kim 2019). In reality, the burden of overcoming these information problems often falls onto the entrepreneurs, many of whom do not possess the necessary knowledge and skills.

The proliferation of data and AI technologies has enabled low-cost, easily-scalable, and highly-customizable entrepreneur training on digital platforms, which has quickly become popular.[3] Most prior studies focus on offline small-and-medium businesses (SMB) training, with mixed and limited results (McKenzie and Woodruff 2014; McKenzie forthcoming). High costs often severely limit sample size and necessitate sample selection, which combine with the inherent noisiness of entrepreneurial success to hinder statistical power. They also preclude tailored training (e.g. consulting) in favor of lectures among new businesses.[4] Also, we have little evidence suggesting that consumers benefit from these programs. McKenzie and Puerto (forthcoming) is a notable exception, but consumer gain is still inferred from market expansion without accounting for prices and consumer choice.

[1] Stats from Marketplace Pulse.

[2] Data from Socialblade Only new creators with more than 5 subscribers are counted.

[3] Examples of entrepreneur training programs: eBay, Amazon, Youtube, and Airbnb.

[4] See McKenzie and Woodruff (2014). Bloom, Eifert, Mahajan, McKenzie, and Roberts (2013) and Bruhn, Karlan, and Schoar (2018) both focus on established plants and above medium sized businesses.

In this paper, we ask three questions: how large is the growth barrier facing new businesses on e-commerce platforms? Can AI-facilitated entrepreneur training reduce the barrier and facilitate new-business growth? How does such training affect consumers?

Our analysis relies on massive amount of new data, randomized roll-out of an innovative training program, and an empirical model of consumer choice and platform steering. We first obtain administrative data from a large e-commerce platform, and match consumers' search, browsing, purchasing, and product return records with sellers' product and operation information. This allow us to detail consideration sets using search sessions as well as the associated search characteristics such as calendar day and search keywords. We can also track actual repeated interactions between a consumer-seller pair.

We also implement a randomized controlled experiment among new sellers on the platform. It features an entrepreneur training program that is customized to each seller. Facilitated by an AI algorithm with negligible marginal cost, it automatically analyzes each seller's product and operating statistics in real-time and offers training on highly actionable business strategies. The program is rolled out to 25% of all new sellers, touching over two million businesses within a year. 24% of those sellers have participated within the first nine months after they become eligible.

Lastly, we develop an empirical model to recover consumer preference and compute consumer welfare. It is estimated on realized consumer choice in a random subsample of our consideration-set data. Compared to our reduced-form analysis, we are able to permit rich heterogeneity in consumers' preference for or their sensitivity to prices and product variety offered by sellers as well as their sectoral preferences. We also use platform-rule enforcement shocks to instrument for product variety and prices. We can thus accurately infer sellers' unobserved quality relative to sector average.

We find three main results. First, the training program boosts new seller growth, and it does so primarily by enhancing their ability to reach more consumers. Using a seller panel, we find that sellers in the treatment group see 1.7% higher sales on average (ITT), supporting an ATT of 6.6%. On the extensive margin, treated sellers are more likely to make at least one sale in the first month of operation, but no effect is detected on long-term survival. This is largely driven by enhanced marketing activities: both the amount of paid ads and the proportion of consumers sourced from these ads increase by increased by 2.9% and 3.3% (ITT), respectively. Although we cannot detect statistically significant effect on overall conversion rate, consumer purchase probability conditional on receiving customer service increased by 1.4% (ITT). Lastly, there is no change in pricing levels or product listings. Since our training program has close-to-zero marginal cost, it therefore delivers high return-on-investment for the platform and the sellers.

Second, we document a growth barrier that new sellers face in the competition for consumer attention. On average, a new seller must achieve 5% higher conversion rate to obtain

the same consumer exposure as a similar incumbent. To effectively control for seller and consumer heterogeneity, we construct a consideration sets based on search sessions, in which the base unit is a consumer-seller-pair that appears in a set. This allows us to show that conditional on receiving consumer exposure, new sellers have significantly higher conversion rates, controlling for *consumer fixed effects* and *intent-effort-date interacted fixed effects*. To construct the latter fixed effects, we measure (search) intent with search keywords and (search) effort with the size of consideration set.[5] To our knowledge, this level of granularity has never been achieved in prior studies. Our demand estimates quantifies the effect of the growth barriers on realized sales and consumer purchase choices. Among sellers with non-zero sales and thus market share, the average new seller provides 23% higher-quality product and service than a comparable incumbent, equivalent to a 65-cent consumer gain.

Third, training generates a market-expansion effect. Comparing two search sessions (consideration sets) done by the same consumer, each with the exact search keyword and search effort, the consumer is more likely to make a purchase if new sellers are part of the consideration set. We also verify that the training program shrinks this conversion-gap between new sellers and incumbents. Combining the three sets of results above, we can attribute the training program's success in boosting new-seller growth primarily to closing the growth barriers. This means that reducing the growth barriers will, at least on the margin, increase new sellers' consumer reach and expand the market.

Our main counterfactual analysis quantifies the welfare impact of reducing the growth barriers through the training program. Specifically, we impose the same growth barrier that the control-group sellers face onto the treatment group, equalizing the all new sellers' likelihood to appear in consideration sets. Overall, only 4.3% of the sellers are new, so that trained new sellers make up only 0.25% of all sellers. However, removing the training program from such a small portion of the market is enough to reduce total revenue reduces by 0.05%, with consumer welfare dropping by 0.07%. In particular, treated sellers suffer 7.7% drop in market share, obtaining a smaller share of a smaller pie.[6]

Related Literature Our study contributes to the empirical literature on firm dynamics and entry barriers. For offline firms, previous studies on entry barriers have focused on difficulties in initial market access (Atkin, Khandelwal, and Osman 2017), slow customer accumulation (Foster, Haltiwanger, and Syverson 2016; Piveteau 2016), and the uncertainty in learning (Arkolakis, Papageorgiou, and Timoshenko 2018; Berman, Rebeyrol, and Vicard 2019). We are among the first to document firm dynamics in online platforms. Closely related is Bai, Chen, Liu, and Xu (2020), which find that SMB exporters face large demand-side search frictions. We bring in consumer-side data to extend their study by quantifying

[5]For robustness we check that consumers are no more likely to request returns or refunds when they purchase from new sellers versus incumbents, while they are as likely to make repeat purchases.

[6]Percentage point are hidden at data provider's request.

a resulting growth barrier that prevents new sellers from reaching consumers as well as its welfare impact. We also experimentally show that a cost-less training program can be effective in lifting that growth barrier, reinforcing the role of information frictions. [7]

Second, we contribute to an extensive literature on business training for SMEs in developing countries, predominately offline and generating mixed results on SMB growth (McKenzie and Woodruff 2014). Online training combined with data and AI technologies solve many well-known big issues of offline training: high marginal cost that prevent customization and scaling, in particular (McKenzie forthcoming). Customization has been shown to be especially valuable (Bloom, Eifert, Mahajan, McKenzie, and Roberts 2013; Bruhn, Karlan, and Schoar 2018). We also find that better marketing is essential for new-business growth online, which echoes Anderson, Chandy, and Zia (2018), Gourio and Rudanko (2014), and Fitzgerald, Haller, and Yedid-Levi (2016). We also add to a growing body of empirical studies featuring "experimentation at scale," (Muralidharan and Niehaus 2017) with a particular focus on evaluating spillover and market-expansion effects. Instead of zooming out and varying treatment intensity across markets (Calderon, Cunha, and Giorgi 2020; McKenzie and Puerto forthcoming), we zoom in to the consideration-set level, and provide a new method based on dissecting consumer choices.

Third, we speak to the growing literature on platform strategy and online marketplaces (Hagiu 2014; Eisenmann, Parker, and Van Alstyne 2006; Cusumano 2010; Parker, Van Alstyne, and Choudary 2016). Our main contribution the documentation of firm dynamics, a growth barrier, and the effectiveness of a training program that is essentially cost-less. In contrast, most existing work focuses on specific economic mechanisms: the design of the reputation and feedback system (Tadelis 2016; Milgrom and Tadelis 2018), pricing (Dinerstein, Einav, Levin, and Sundaresan 2018), auctions (Agarwal, Athey, and Yang 2009; Backus, Blake, Larsen, and Tadelis 2020), and search ranking (Ghose and Yang 2009; Ursu 2018; Horton 2017; Barach, Golden, and Horton 2020).

The rest of the paper is organized as follows. In Section 2, we describe the e-commerce platform, new-seller growth, our data, the training intervention, and the experimental design. Section 3 presents the average impact of the training on new sellers. In Section 4, we analyze consumers' consideration sets, and show that training has a market-expansion effect, and that it does so by mitigating a growth barrier. In Section 5, we present and estimate a demand model to account for consumer heterogeneity, quantify the effect of the growth barrier, and calculate the welfare impact of reducing the barrier with the training program.

[7] 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, Einav, Levin, and Sundaresan (2018) (eBay), Horton (2014) (labor market), Ursu (2018) (Expedia) and Chen and Wu (2020) (AliExpress).

2 Background and Data

In this section, we introduce the setting of our study, our administrative data, and our main experimental intervention. Section 2.1 provides background information on the e-commerce platform, focusing on key platform rules for new sellers and anecdotal evidence of a growth barrier they face. Section 2.2 introduces the AI-facilitated entrepreneur training program. Lastly, Section 2.3 details our experimental design and the dissemination of the training program.

2.1 E-commerce: the Platform and the Sellers

We partner with a leading e-commerce platform in China that hosts millions of active sellers and hundreds of millions of consumers. Sales facilitated by the platform represent a sizable share of all Chinese retail. Product variety encompasses all major retail sectors, including clothing, cosmetics, home appliances, consumer electronics, and food.

Most sellers on the platform are third-party retailers who source their products from manufacturers and/or wholesalers. The platform directly sells to consumers only in a few sectors, notably fresh groceries and financial services. Our analysis excludes all such sectors and thus feature those in which the platform only hosts—and does not compete with—sellers operating on it. In our context, the platform thus profits only via advertisement auctions, sales commissions from large sellers, and paid supplementary services such as advanced data analytics (Jin and Sun 2019).

Entry The platform facilitates easy entry by new businesses, particularly for retailers that have already procured products. The registration process is close-to-cost-less: an individual entrepreneur can directly register as a sole proprietor seller on the platform with a national ID or a formal business registry number. The platform also offers intuitive drag-and-drop user interface to help sellers easily showcase products and build virtual storefronts. The platform prevents abuse and deters bots by charging a small deposit (returned to sellers after two weeks),[8] and with the help of automated identity verification against the national digital identity database. Moreover, there are several requirements for product listings. For example, high-resolution product pictures are required, and they cannot be duplicated or partially-duplicated with other product images on the platform. In most cases, such requirements are enforced with AI algorithms so as to not delay the entry process. In 2019, about 25,000 new sellers registered every day.

Platform matching and seller strategy Traffic is the lifeblood of any online business. For e-commerce sellers, this is measured by the amount of unique visitors in a given period

[8]This deposit is charged before posting any products and ranges from $0 to $700 across product categories. If sellers decide not to put down the deposits, they can still post products on the platform, but their products will get much lower rankings in search results and will not be promoted in other channels.

of time. Almost all traffic to new sellers are intermediated by the platform:[9] Consumers arrive at the platform and search for a desired product; the platform then dissects the search query in to keywords, and uses sophisticated AI algorithms and position auction to rank a set of related products and sellers. Correspondingly, sellers can boost ranking by accumulating past sales and positive reviews, refining product descriptions, offering price discounts, maintaining fast customer service, and marketing, which includes paying for ads through position auction.[10] Seller interviews suggest that marketing and customer services tend to make up the largest component of G&A cost, especially among large sellers.

New-seller growth and industry dynamics While seller entry is easy, growth is hard. Almost all sellers that have posted products receive some initial traffic from the platform in the first month, with 50% of them completing a sale. For the average cohort of new sellers (with at least one product listing), the first month sees 0.05% and 1.7% of them ranked among the top 1% and 10% by traffic in their respective retail sectors. However, at the end of the first year, growth among these new sellers becomes more unbalanced and concentrated at the top. While 0.08% of them are now placed in the top 1%, only 1.3% of are placed in the top 10%.

Administrative data Our main data source is the administrative data that the platform collects on the sellers and consumers' it hosts. On the seller side, we have data on performance, strategies, characteristics and participation in the training. In addition to the usual performance measures such as revenues, traffic and conversion rates[11], we also observe quality metrics that include ratings, frequencies of refunds, requests and rule violations. Besides outcomes, we collect various proxies for sellers' product, pricing, and marketing strategies and have information about whether or not the sellers registered as formal firms and their geographic locations. Lastly, we obtain detailed records on sellers' participation in the training program for the entire history of task take-up and completion.

On the consumer side, we obtain rich records of consumers' searching, browsing and purchasing behaviors that allow us to reconstruct consumers' actual consideration sets and explore the variations in the compositions of these sets. For an individual consumer on a given day, we observe which sets of products (sellers) she visited and which ones she bought after each keyword search on the platform. We track the quality of the purchase with subsequent requests for returns, refunds and repeat purchase. The consumer-product-keyword level results thus allow us to directly link individual consumer with specific sellers to evaluate the interactions. The underlying data also allows us to construct

[9]Typically, only an established seller can source a significant amount of traffic from outside the platform with off-platform advertising and marketing.

[10]For the position auction, sellers submit pay-per-click bids, and the top few bidders will be placed in between natural search results, in addition to their natural ranking.

[11]Without access to costs, we are unable to measure profits.

features that characterize consumers' past experience and preferences.

2.2 An AI-Facilitated Entrepreneur Training Program

In order to facilitate new-seller growth, we collaborate with the platform to develop and implement an AI-facilitated entrepreneur training program. The key hypothesis is that new sellers suffer from information friction about platform operations and effective online business strategies. This can be captured by administrative data of their actual strategic choices, or lack therefore. We mainly focus on marketing and customer service, which are critical in attracting traffic and improving sales conversion (conditional on customer visit).[12]

Content of training We partner with several third-party e-commerce service providers to design intuitive and highly-actionable training materials. In particular, a training program is organized as a sequence of individual tasks, each introducing a specific business strategy or highlighting an area of improvement (and the corresponding strategies to implement). Each task consists of tutorials, Q&A forums and webinars to deliver recommendations. Tasks also have varying completion time based on their difficulty level, ranging from three to thirty days.

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



[12] We do not focus on pricing decisions because within-seller pricing variation is largely driven by platform-wide sales, and is thus undifferentiated among new and established sellers and across sellers with various growth profiles (Jin and Sun 2019). However, our empirical analyses account for the effect of prices on competition, consumer choice, and traffic allocation.

For a new seller, the training tasks cover three major areas: basic setup, marketing strategy, and customer service[13]. The first type of task help sellers improve product postings and virtual store setup and avoid violating platform rules. For example, product postings can suffer from incomplete descriptions or uninformative pictures. A corresponding task contains case studies of good and bad product pictures, a step-by-step guide to set product labels, and a checklist to refine product descriptions[14].

The second type of tasks teaches marketing techniques to attract visitors and increase store traffic. For example, search-engine optimization (SEO) is critical in obtaining organic traffic. One of the corresponding tasks contains detailed information on how product titles relate to search rankings, and sector-specific case studies that highlight the importance of avoiding irrelevant or overly popular keywords. Another series of the tasks introduce how the platform auction off some of the most attractive search positions in pay-per-click position auctions. For example, some tasks explain auction rules and regulations, while others focuses on practical guide on how to submit bids or on time-of-day-specific discounts that are particularly important to new sellers, etc.

The third type of tasks is about improving the quality and efficiency of customer service. An example is a comprehensive introduction of AI-based customer bots tailored for high-growth new sellers. It outlines simple procedures to set up automated responses as well as the appropriate use cases such as responding to inquiries about business hours and delivery status. This series of tasks focuses on tapping into these resources to grow customer service capabilities at low cost. Other service-related tasks teach the etiquette when interacting with customers, with detailed suggestions on how to address customers' concerns effectively to improve conversion.

AI customization and participation We also develop a machine-learning algorithm that utilizes a new seller's actual operational and product data to determine the relevance of each available task and the sequence with which the tasks should be assigned. In addition, this process is dynamic, so as to ensure that any task shown to a trainee is appropriate given the current state of that trainee. Moreover, each task is associated with specific performance metrics along with the corresponding triggering and completion conditions. For example, a new seller could trigger the task "attract more visitors to your store" if the number of visitors she had over the past 30 days is lower than 60 percent of sellers in her sector. When this seller successfully implemented the strategies in the tutorial or when it has successfully surpassed the visitor threshold, the task will be marked as completed au-

[13] Appendix table A15 provides a list of tasks with explicit contents and classifications.

[14] There are additional complexities of the product management system. For example, each posting requires sellers to describe the characteristics of the products in great detail (setting product labels), which is a direct input of the platform's search algorithms. There are also many tricks that can help boost exposure of the products. For example, to get the product promoted in one of the major recommendation channel, at least one of the pictures should have a white background.

tomatically.[15] Taken together, at given time, each new seller is offered a training program that is customized to its needs.

Sellers can choose to take up one or more training tasks at any given time from the list of available ones (figure 1 panel B). They can then access detailed tutorials, and can also reach out to the training provider directly via email[??]. Some tasks include live-streamed webinars to facilitate real-time interactions and Q&A. If a seller fails to complete a task in the designated time frame, it can always make additional attempts later.

Distribution and Scope The training program is offered at no cost to new sellers. It is prominently featured on the platform's official (and only) seller-facing portal as a widget on the front page (figure 1 panel A)[16], which guarantees effective exposure to sellers.

The program was launched on May 6, 2019 as part of our experiment outlined below. Over two million new sellers have accessed the training as of June 2020. Despite such prominence, it is a standalone program that is independent from other operations of the platform. Training participation and performance do not affect how the platform matches consumers with sellers in the searches.

2.3 Experimental Design and Implementation

We design and implement the training intervention as a randomized controlled experiment by randomizing access to the training program among new sellers. The access is assigned on the day when new sellers complete the registration[17]. 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[18]. Moreover, the time-varying, dynamic matching between sellers and tasks forbid us from randomizing at the task level.

The experiment started on May 6, 2019 and has been ongoing since[19]. In 2019, about 25,000 new sellers registered every day. Figure 2 summarizes the roll-out of the training. New sellers are first randomly selected into the experimental sample. Among those in the sample, some are randomly selected to get access to the training (in yellow). From sellers' perspective, there are no differences in experience to be in the control group (in green) or to be out-of-sample (in purple). Our empirical analysis focuses on new sellers registered

[15]For each completed task, sellers earn short-term free accesses to certain supplementary services. These supplementary services support routine online business operations. As an example, sellers can access a program that allows sellers to print shipping labels in batches. The monetary value of the rewards ranges from \$5 to \$10. Sellers are unaware of the rewards before taking up their first tasks.

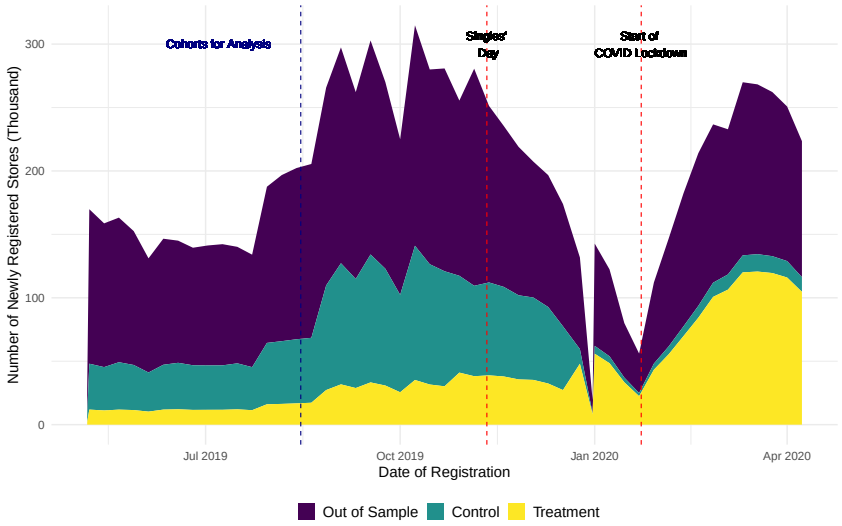
[16]The app is available in all major mobile operating system and has a web version.

[17]Vast majority of the sellers complete the registration process in a single session within a day. Notice that posting a product is *not* required to complete the registration. In fact, as many as 30% of registered sellers never post any products.

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

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

Figure 2: Experimental Sample and Treatment Assignment Over Time



between May 6 and August 15, 2019, part of the first phase of the formal roll-out of the training intervention. Among the new sellers in the experimental sample, we randomly selected 25% of new sellers to access the training. Later on, we expanded the share of new sellers in the experimental sample to access the training to 35% on October 29 and to 90% on December 26. By June 2020, over two million new sellers have gained access to the training program.

2.4 Training Take-up

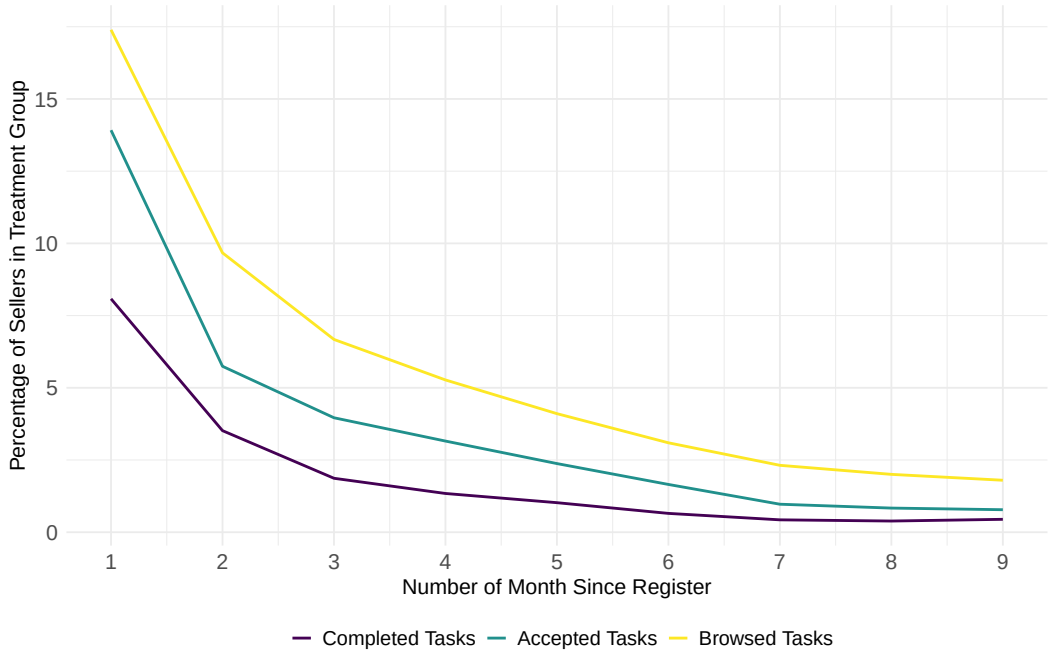
Among sellers in the treatment group, 44.6% browsed the training during the first month of entry. Within the first nine months after registration, 24.1% of treated sellers took up at least one task and 12.6% completed at least one task. The randomization was successful since none of the sellers in the control group took up any tasks. This adoption rate is typical for online products, but lower than many offline training programs (McKenzie and Woodruff 2014)[20].

Sellers are most active in training participation during the first month of entry, partially because the content of the training is most relevant for the fresh entrants. Altogether, 49.9% of tasks were taken-up during the first month of entry. During this time, sellers are more likely to select tasks related to setup and customer acquisition than those on customer management. New sellers may not find the training particularly useful to help them go through the process of posting products. Among sellers in the treatment group who

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

eventually posted products, only 7% took up the relevant tasks before posting. However, tasks related to store setup do have a higher completion rate at 58.0%, comparing to the average rate at 37.5%[21].

Figure 3: Training Participation and Retention



Although sellers have access to the training for at least six months after the registration, the retention rate declined relatively quickly over time. 13.9% of treated sellers took up tasks during the first month, but only 4.0% of sellers continued to do so in the third month[22]. Conditional on taking up any tasks in the previous month, around 23% of sellers took up more tasks in the following month. Figure 3 shows the share of sellers who browsed, took up and completed the tasks in the subsequent months following the entry.

There are sizable variations in the degree of involvement. Among sellers who participated in the training, 29.8% only took up one task, yet about 9.6% of sellers took up more than 10 tasks. Overall, the median participant took up 3 tasks. Figure A1 shows the distribution of the number of tasks seller took up. In terms of completion, 61.2% of participants completed at least one task, while the average participants completed 1.63 tasks. The 5.5% most active participants completed more than 6 tasks. Overall, sellers who took up more tasks also have higher completion rates. There are also sizable variations in com-

[21]The higher completion rates might be due to the fact that these tasks are easier to complete to begin with. Most of the tasks in this category ask sellers to complete certain actions, rather than to out-perform their peers.

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

pletion rates across different tasks. The basic setup tasks have an average completion rate of 42.9%, but the tasks that focus on improving customer service quality only have average completion rate of 35.2%[23].

Columns 4 to 6 of table A1 summarize characteristics of the treated sellers by whether or not they participate in the training. Compared to sellers in the treatment group who did not participate, sellers who took up tasks are slightly more likely to be registered as firms but among those registered as individuals, the participating sellers are less likely to be female[24]. Sellers from more economically developed southern coastal provinces are more likely to participate and take up more tasks conditional on participation[25]. Sellers from the southern provinces contribute more to total sales on the platform than sellers from anywhere else. We hypothesize that sellers who post products on the very first day of entry are better prepared. These early-movers are more likely to become training participants, but conditional on participation, they do not take up more tasks.

3 Training Boosts New Seller Growth

In this section, we construct a seller-level panel data and evaluate the impact of training on seller growth in our RCT. In Section 3.1, we detail our sampling strategy, which is necessary due to the massive size of the platform and hence our raw data. We are then able to directly estimate treatment effects in Section 3.2. Overall, we provide direct evidence that training boosts new-seller growth, both in traffic and in realized revenue. This is largely driven by enhanced marketing and customer service effort.

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 participation in the training[26]. Our sample covers a subset of new sellers registered between May 6 and August 15, 2019. On average, 22,230 new sellers registered on the platform per day during the sample period.

[23]Tasks focusing on marketing, customer services and customer management have completion rate of 39.7%, 22.3% and surprisingly 3.4% respectively.

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

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

[26]Similar to other studies (Zhang, Dai, Dong, Wu, Guo, and Liu 2019) on online businesses, our data only includes activities that are observable online. Specifically, we do not have data on costs, including marketing expenditures. 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 suggests that the vast majority of the sellers have a minimal offline presence and operate solely online.

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 profile to be included in the final sample[27]. Since sellers will not find out whether they have access to the training program before the first login to the Seller's Portal, the login requirement does not induce selection. However, the training could affect the frequency of subsequent logins[28]. 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 A1). All the characteristics are balanced across the treatment and control except for sellers' registration type, where treated sellers are more likely to be firms but the difference is small[29]. Among sellers registered as individuals, 45.3% are females. Gender distribution is balanced, so are the locations at both province and city level.

We construct a balanced monthly panel with all the sellers in the final sample. The panel spans the subsequent nine months following the registration, where we define each month as a 30-days period relative to the date of entry. This entry day is also the day of treatment assignment. We collect measures on sellers' revenues, numbers of visitors and conversion rates[30]. We also collect quality measures, including sellers' customer ratings (evaluated on the accuracy of the product description, quality of customer service and logistics), the likelihood for consumers to request refunds or returns and the frequency of rule violations[31]. On sellers' strategies, we observe their product offerings, pricing level, involvement in marketing and provision of customer service. We do not observe actual spending on marketing. Instead, we observe proxies for sellers' engagement, such as the number of products participating in pay-per-click ads. For sellers with treatment access, we obtain their entire history of interactions with the training program. We focus on number of tasks take-up and completion in the analysis[32].

In the final sample, 40.6% of registered sellers have never posted any products to sell. For sellers who posted products or earned revenues, we obtain their affiliated sectors[33]. For

[27]Sellers would have their basic information and locations recorded in the profile. We impose these requirements to screen out inactive new entrants.

[28]We empirically test the treatment effect on login but do not find any significant differences.

[29]26.4% of sellers in the control group are registered as firms. The share of firms in the treatment group is 26.8%. Our main results continuous to hold controlling for sellers' type.

[30]Conversion rate is the defined as the share of visitors who make purchases within a designated period. The conversion rate we use in the analysis is in the same suit as that the platform uses the conversion rate to evaluate its search and recommendation algorithm's efficiency.

[31]Platforms impose many rules to regulate the sellers. Typical rule violations include using inappropriate words in the product descriptions, illegal use of intellectual properties, and placing fake orders.

[32]As described in section **??**, the assignment of tasks is individualized with temporal variations. We do not keep track of the daily updated actual task assignment.

[33]Sellers' sectors are determined by the 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 is labeled as selling

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[34]. 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 leave these variables as missing. Distribution of the outcome variables such as the number of visitors and revenues is extremely skewed. For the main analysis, we convert these variables to log scale[35].

3.2 Impact of the Training

Overall Impact We first evaluate the overall intent-to-treat (ITT) effect of the training 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 1 presents the estimated results on sellers' performance. Access to the training leads to a 5% increase in the likelihood of earning revenues (column 2) as well as a 1.7% increase in revenues earned (column 1)[36]. Since revenues are zero for 60.5% of seller-month pairs, the unconditional average revenue is close to zero. Restricting the sample to seller-month pairs with positive revenues (column 3), sellers in the treatment group earn 2.6% higher monthly revenues. The magnitude of the treatment effect on revenues is positive

second-hand products, which are treated differently in the search results. We group sellers without sectors, sellers who sell 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. The results do not change if we use sellers' most frequently affiliated sectors.

[34]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 the 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. We do not exclude sellers without product posting in the main analysis because encouraging sellers to post products is an important part of the training. Hence the decision to post products could be affected by the training. We do not find a significant difference in the likelihood of posting products in the subsequent months, but we cannot rule out the possible treatment effect.

[35]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 measures before taking logs. We also use inverse hyperbolic sine transformation for the main outcomes and reach a similar conclusions.

[36]Table A2 presents the treatment effect on revenues in levels with different winsorization thresholds. To address the concern of using log scale with many zeros, we also use inverse hyperbolic transformation for revenues (column 4) and find similar results. 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 contents of the training, we do not expect this training could have meaningful impact on the top sellers.

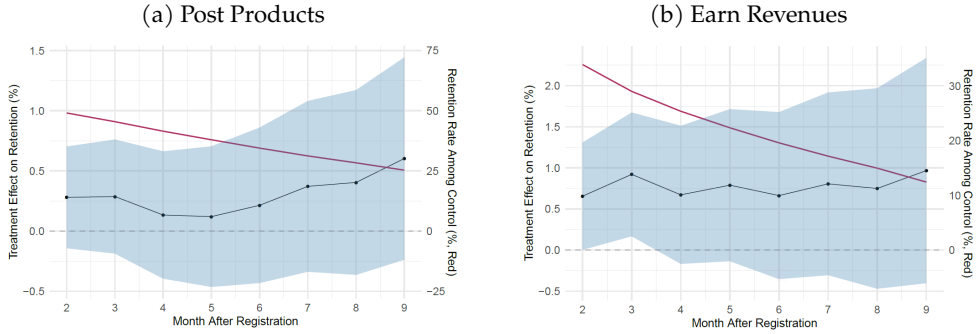
but small. Over the nine months, treated sellers earned $1.8 million higher total revenues. Assuming the treatment effect is of a similar magnitude for all cohorts of new sellers, all two million treated sellers combined would earn about $4.7 million extra revenues.

Training helps new sellers gain more consumer exposure as they attract more visitors to their sites. When decomposing the revenue increases, we find that 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). Conditional 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 in treated sellers' conversion rates (column 8). We use consumer-side data to explore more nuance impact on conversion in section 4.

Table A8 summarizes the impact of training on sellers' observed quality metrics. We do not find that the treatment significantly improves customer ratings. The two groups of sellers also have similar likelihood of requests for refunds and complaints, shares of positive reviews, and frequencies of violating the platform's rules. Taken together, sellers with training access weakly outperform control sellers for the observed quality.

Focusing on sellers who have participated in the training, we use the random assignment of the training access as an instrument for the actual participation. The first stage outcome is an indicator for taking up any tasks during the sample period. Column 1 of table A5 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, but do not have significantly higher 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 outperform non-participants on all performance measures (table A7).

Figure 4: Treatment Effects on Retention



Notes: The black connected lines (left-scale) present ITT estimates on retention using different metrics. The dots are estimated coefficients of β in Equation (1)

Table 1: Overall Treatment Effects on Sellers' Performance

	Dependent variable:							
	Log Revenues	Any Revenues	Log 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 outcomes for all sellers in the new seller sample. 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 the share of visitors that make a purchase within a week. Column 3 restricts the sample to sellers that earn non-zero revenue during the month. Column 5, 7 and 8 include only sellers that have visitors during the month. All regressions include cohort, relative month and main sector fixed effect as described in equation 1. Dependent variable means are calculated with sellers in the control group across all months. Standard errors are clustered by seller. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

We analyze temporal variations of the treatment effects in appendix A.1 and conclude that the temporal variations are small. In particular, the training does not have lasting impact on sellers' retention. Figure 4 presents the treatment effect on retention measured by having products posted and earning revenues over time. Treated sellers are no more likely to have products posted throughout. They are more likely to earn revenues but only during the initial months. The experimental design limits the baseline heterogeneity we could capture to basic types, locations and actions taken on the first day. We discuss the heterogeneous treatment effects in details in appendix A.2. In a nutshell, we do not find a significant difference by sellers' registration type or level of preparedness[37]. Instead, the offline business environment could play a role. Sellers from less-developed regions are less likely to participate in the training, while the training is not as useful for these sellers.

Sellers' Strategies Next, we discuss what strategies that new sellers undertake could lead relate to changes in the observed increase in revenues and the number of visitors. We focus on 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 2 presents estimated coefficients β in specification 1 on the treatment indicator, where each cell corresponds to a separate regression on a specific outcome variable. Although many tasks focus on setting up online stores, the training has limited impact on incentivizing market participation in terms of posting products or putting down security deposits. Here, 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 A3 columns 1 and 2). Similarly, treated sellers are no more likely to put down security deposits (table 2). As expected, sellers do not behave differently in terms of the number of products offered, the likelihood of moving into different sectors or setting different prices. These strategies are not explicitly discussed in the training.

Treated sellers are more likely to follow the platform's recommendations to adopt free supplementary tools that help improve the quality of customer service[38]. Specifically, we find that treated sellers have a slightly shorter average response time when consumers make inquiries and have higher conversion rates among consumers who have made in-

[37] We measure the level of preparedness by whether or not sellers post products on the first day of registration. Sellers who are ready to post products on the very first day are arguably more prepared and more proficient.

[38] The supplementary tools that the training introduced focus on improving customer service quality and are not the same as those given out as rewards for completing the training tasks

Table 2: Treatment Effect on Sellers' Strategies

Variable	Treatment (%)	Dep Var Mean	Variable	Treatment (%)	Dep Var Mean
Market Participation			**Marketing**		
Post Products	0.29	0.34	Paid Ads (Freq)	2.86**	0.07
	(0.29)			(1.43)	
Paid Deposits	0.32	0.31	% Visitors from Ads	3.33**	0.03
	(0.32)			(1.33)	
			Limited Time Sales	25.00*	0.0008
Service				(12.50)	
Active Time (min)	-0.41	19.17			
	(0.67)		**Pricing**		
Reply Time (min)	-0.29*	389.02	Avg. Price Per Buyer	0.11	4.39
	(0.15)			(0.14)	
Conversion Rate	1.40***	0.1	Avg. Price Per Product	0.008	3.83
	(0.50)			(0.026)	

Notes: Table presents estimated coefficients β on treatment assignment dummy with specification 1. Standard errors are 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 a seller' account is active and can answer customer inquiries. Reply time is average time customers waited to hear responses from sellers when making inquiries measured in minutes. Conversion rates are measured as share of consumers making purchases among those who have contacted the customer service and made inquiries. Paid ads (product counts) is the number of products participating in pay-per-click ads. Paid ads (traffic share) is share of consumers visiting sellers' sites from paid channels (including both pay-per-click ads and other channels). Promotion is the number of times sellers participate in the limited time promotional events that the platform regularly organizes. Average price per products and average price per buyers are are seller-level prices measured in log scale. The prices are transaction price, rather than the listed prices. 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.

quiries (table 2 section on customer service). These results are driven by treated sellers' higher likelihood of adopting the AI-backed customer assistant to help answer consumers' basic questions[39]. Although treated new sellers do not have significantly higher customer ratings of their service quality, improved customer service quality could still contribute to treated new sellers' revenue increases as indicated by higher purchase probabilities among consumers who make inquiries.

Training could incentivize sellers to invest more efforts on marketing (table 2 section on marketing). 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 a higher share of visitors coming from the paid channels[40]. In addition to advertising, treated sellers are also more likely to participate in the limited-time promotional events that the platform regularly organizes[41]. The products on sale get preferential treatment in the search rankings and additional exposure in channels other than direct search. More active partic-

[39]Currently we do not have data on sellers' actual subscriptions to the AI assistants and other supplementary service.

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

[41]The classic sales events the platform regularly offers are product specific limited-time discounts. Sellers select the set of products to participate and submit applications to the platform to be included.

ipation in marketing could involve higher costs for the treated sellers. However, without data on actual spending on marketing, we are unable to claim that the training definitively increases sellers' profits.

Marketing is an indispensable part of online business operations, and attracting visitors is the key to success. Previous experimental study found that consumers interpret costly advertising as signals of high quality (Sahni and Nair 2020). The training intervention that either improves sellers' marketing skills or raises sellers' awareness of marketing helps new sellers attract more visitors potentially at the expenses of incumbents and non-treated new sellers[42]. Such reallocation could be welfare enhancing for consumers if they interact more often with a higher quality group of new sellers who are willing to invest in marketing. We turn to the impact of the training on consumers in the next section.

4 Beyond Testing Information Friction: the Growth Barrier and Consumer Welfare

In the last section, we have shown that our entrepreneur training program produces sizable growth impact on new sellers. This validates our hypothesis that information frictions exist: many new sellers do not have a comprehensive grasp on platform operations and effective online business strategies, particularly with respect to marketing and customer service. In this section, we go beyond the detection of information friction—we quantitatively identify a growth barrier that new sellers face in the competition for consumer attention, and the degree to which the training program has lowered it. We also show that reducing this barrier increases consumer welfare.

To do so, we merge our seller dataset with detailed record of consumers' search and browsing behavior as well as their purchase and (product) return decisions (Section 4.1). Specifically, in the following analyses, we are able to control for *consumer fixed effects* and *intent-effort-date interacted fixed effects*. To construct the latter fixed effects, we measure (search) intent with search keywords and (search) effort with the size of consideration set. In addition, we also control for sellers' past price level, ratings, and product variety as well as basic consumer characteristics. To our knowledge, this level of granularity has never been achieved in prior studies.

In Section 4.2, we examine consideration sets with at least one new seller. We show that, conditional on receiving consumer exposure, new sellers have 5% higher conversion rates, which demonstrates a growth barrier they face. In Section 4.3, we expand our attention to

[42]Since we do not observe sellers' operating or marketing costs including their spending on the platform, 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 impact on sellers' available financial resources to invest in marketing, even though sellers might shift investments from other channels to marketing.

all consideration sets. We show that the training program produced a market-expansion effect, increasing consumer welfare, which also implies that the growth barrier identified above hurts welfare. Specifically, comparing consideration sets with and without new sellers, we find that including a new seller to a consumer's consideration set—regardless of the seller's treatment status—increases the consumer's purchase likelihood by 1.9%.

In Section 4.4, we conduct two robustness checks. First, we show that consumers are no more likely to request returns or refunds when they purchase from new sellers versus incumbents, while they are as likely to make repeat purchases. Second, adding consider-set fixed effects, we find that the training increased the purchase probability (or conversion rate) of new sellers. This means that the aforementioned market-expansion effect is at least partially coming from increased allocation and market share towards new sellers.

4.1 Consideration-Set Data and Sample Construction

To evaluate how interacting with new sellers affect consumers' experience, we identify consumers with the same interests (searching the same query) and the same search efforts (visiting the same number of sellers) on the same day but visit different sets of sellers. To be precise, we construct a consumer-search session sample using administrative data from the platform in the following steps (illustrated in A4). First, we draw a random sample of new sellers from the experimental sample[43]. Next, we identify a set of consumers who have visited the sampled sellers between August 1 and December 31, 2019. We also obtain the search keywords the consumers used to find these sellers. Then, we find another set of consumers who searched the same keywords on the same day and visited the same number of sellers as the previously identified consumers did but only visited incumbents[44]. As mentioned, sellers that each consumer visited are not randomly selected because both search rankings and consumers' browsing behaviors are endogenous. However, without explicit efforts, it is almost impossible for consumers to specifically look for new sellers when searching for products[45]. Therefore, our empirical strategy exploits variations in the

[43]We have to draw a sub-sample mainly due to computational constraints. All sellers in the experimental sample combined attract more than 10 million visitors on a typical day.

[44]We limit the number of matched consumers 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 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 about 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 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.

[45]A consumer can only determine whether the seller is newly registered if they visit the seller's site and click on a specific section about the store's necessary information. In the search results, besides prices, product titles, and category-specific labels, the displayed information also include the cumulative number of orders and sellers' locations. The cumulative number of orders imperfectly correlate with sellers' age. Consumers can search specifically for new stores by including relevant keywords in the search query, but the search results

platform's search algorithm, which determines the pool of sellers that consumers interact with. For the 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 after the event. Final sample is at consumer-search session level. A search session is a search query-search efforts-date combination[46].

From the identified consumer-seller-search session combinations, we then construct dataset at different granularity for the subsequent analyses. From the sample construction process, we obtain the most granular level data at consumer-seller pair level, where each pair belongs to a specific consideration set (i.e. search keyword-date-effort combination, referred to as "consumer-seller pair sample").[47] The consumer-seller pair data therefore forms the basis of the first type of dataset. We restrict the main data to different subpopulation based on specific needs in the following analyses. For each consumer-seller pair, we use whether or not a consumer makes a purchase from the specific seller within a given time period as well as the size of the order as the main outcomes. The main outcome variables are whether or not the consumer makes purchases from the specific seller within a designated period of time. When a consumer makes a purchase, we obtain quality measures on whether or not the consumer requests returns or refunds and whether or not she makes repeat purchases from the same seller in the following month. We obtain sellers' pricing, the number of products offered, and their ratings as the main control variables. For the consumers, we obtain their search and purchase history around the time of the session, which allows us to summarize consumers' characteristics and preferences.

For the second type of data, we aggregate the previous master dataset to consumer-search session level so that each observation is a consideration set that forms the basis of consumers' purchase decisions ("consideration set sample"). We include the same set of the seller- as well as consumer- side controls as in the first type of data, aggregating those measures at consideration set level. The master dataset covers the exact same group of consumers and their consideration sets as described in the previous paragraphs.

The final sample consists of 1,381,273 consideration sets (consumer-search sessions) spanning 153 days in the second half of 2019. Table A11 summarizes the main variables with the consideration set level data. The sample consists of 515,748 consumers. On average, each consumer appears in 2.68 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 18.8% of the search sessions, consumers placed an order with some sellers they visited on the same day. Average consumers visited

do not always respect that intention. Moreover, we control the search keyword in the primary analysis.

[46]Search efforts are defined by number of sellers that consumers visited.

[47]We also restrict the sample to the set of consumers who appear in at least two search sessions. Since we control for consumer fixed effect in the main specification, including consumers with one search session does not change the results.

4.89 sellers per search session.

Interacting with new sellers is rare[48]. In our sample, consumers visited treated new sellers in 6.5% of search sessions and control new sellers in 9.2% of sessions[49]. 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 have none or very few visitors. Therefore, the consumers' results do not speak to the average new sellers, but only new sellers in the top end of the distribution whom the consumers can encounter when searching on the platform.

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

4.2 Existence of the Growth Barriers

In this section, to check the existence of growth barriers, we analyze consumers' purchase probabilities when they interact with new sellers in searches. Specifically, we answer the following question: when a consumer visits both new sellers and incumbents in a search session, from whom is she more likely to purchase? We construct a consumer-seller matched pair sample to analyze consumers' choices and the resulted market allocation.

Empirical Strategy To test for the existence of the growth barriers, we restrict the main consumer-seller pair sample to consideration sets that have at least one new seller and run the following specification:

$$Y_{ijs} = \beta^t \mathbf{N}_j + \mathbb{X}_j\gamma + \alpha_i + \alpha_s + \epsilon_{ijs} \tag{2}$$

The specification includes consumer fixed effects α_i and search keyword-effort-date fixed effects α_s to address relevant 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

[48] In section **??**, we analyze the number of visitors *new sellers* attracted and compare between treated and control *new sellers*. In this section, the goal is to compare new sellers to *incumbents*.

[49] As described in the sample construction process, the probability of interacting with new sellers here is calculated based on the sample. It 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.

session specific idiosyncratic variations and only evaluate consumers' choices between sellers in the sets. The coefficient of interest is that on $\mathbf{N}_j$, indicators for control-group sellers. We also include a set of seller level controls $\mathbb{X}_j$ on seller j's pricing, the number of products offered and its ratings.

We use α_s to capture variations across search sessions, where each session s is a search keyword-effort-date combination. Controlling for the search keywords addresses potential differential responses when consumers search for different types of products. For example, consumers may be more willing to explore new options when searching for clothing and snacks than for durable and household supplies. Controlling for the number of sellers that consumers visit in a search session helps alleviate two competing concerns. On the one hand, consumers visiting more sellers are more determined to make purchases as they invest more search efforts. On the other hand, consumers browsing more sellers might also be less satisfied with their previous matches, forcing them to search more intensively and lower their overall probability of purchase. Therefore, the consideration set's size serves as a proxy for the intensity of the search and quality of sellers encountered. Similarly, we control for the 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' idiosyncrasy in their pickiness, purchasing power, experience and familiarity with the platform.

A growth barrier Table 3 presents the results of specification 2. Consumers are nearly 5% more likely to purchase from control-group new sellers on the day of visit (column 1) and spend 1.7% more. Such difference in the conditional conversion rate also points a growth barrier: new sellers need to realize higher purchase probability in order to appear in the same consideration sets as their competing incumbents do.

One concern one may have regarding these results is that it is unclear if the new sellers' higher purchase probabilities come solely at the costs of incumbents (business-stealing), or if they are able to expand the market overall with more demands meet. We directly tests for the existence of market expansion in the next section.

4.3 Market Expansion by Reducing the Growth Barrier

To test for possibility of market expansion, we use the consideration-set sample and ask, 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 a purchase? We analyze search session level purchase probability in order to suppress the market re-allocation that plays a role in the previous section.

Table 3: Consumers' Interactions with New Sellers

	Dependent variable: Purchase		
	Purchase		Spending (Log)
	Same Day	In a Week	Same Day
	(1)	(2)	(3)
Control Seller	0.002***	0.003***	0.008***
	(0.001)	(0.001)	(0.003)
Incumbent Mean	0.04	0.048	0.17
Treatment - Control	0.0018*	0.00071	0.0087*
	(0.00099)	(0.0011)	(0.0046)
Observations	1,262,921	1,262,921	1,262,921
R^2	0.234	0.226	0.232
Adjusted R^2	0.056	0.047	0.054

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 (Equation 2). Outcome variable in columns (1) to (2) are dummies for consumers making a purchase from any seller on the day of visit and within 3 days of visit, respectively. Column (3)'s outcome variable is total payments made by the consumer. * p<.1, ** p<.05%, *** p<.01.

Empirical Strategy We answer this question using the consideration set sample with the following specification:

$$Y_{is} = \beta^t \mathbf{T}_{is} + \beta^c \mathbf{C}_{is} + \mathbb{X}_{is}\gamma + \alpha_s + \alpha_i + \epsilon_{is} \tag{3}$$

for consumer i and search session s. The main variables of interests are $\mathbf{T}_{is}$ and $\mathbf{C}_{is}$, defined as having any treated or control new 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 not subject to selection. In addition, we add control variables $\mathbf{X}_{is}$ that includes pricing level, number of products offered and ratings as well as consumer i's total spending and search intensity in the previous week[50].

Results Table 4 presents the results of specification 3 on indicators for purchase probability and spending. On the day of the visit, compared to visiting a set full of incumbents, a consumer is more likely to purchase if her consideration set has new sellers. Encountering treated or control new sellers in the set increases purchase probability by 1.7% and 2.2%, respectively (column 1). Conditional on consumers' recent behaviors and sellers'

[50]Sellers' ratings are determined by the cumulative number of positive reviews sellers get. Hence they reflect sellers' size more than quality. The price level is measured in the log scale. All the measures are aggregated to the consideration set level by average across sellers appeared in the set.

Table 4: Consumers' Interaction with New Sellers and Purchases

	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	0.004***	0.004***	0.005***	0.014**	0.017**
	(0.001)	(0.002)	(0.002)	(0.007)	(0.007)
Control Seller	0.005***	0.004***	0.005***	0.028***	0.030***
	(0.001)	(0.001)	(0.001)	(0.006)	(0.006)
Incumbent Mean	0.18	0.21	0.22	0.76	0.78
Treatment - Control	-0.0074	0.00004	-0.0065	-0.014*	-0.013
	(0.0018)	(0.0019)	(0.002)	(0.0081)	(0.0086)
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 3. Outcome variable in column (1) to (3) are dummies for consumers purchasing from any seller in the consideration set on the day of visit, within 3 days and within a week of visit, respectively. Spending is total payments made and order size is the size of order before applying discounts. The latter 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-tests for listed coefficients with standard errors. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

characteristics, consumers are 1.9% more likely to make a purchase with one more treated or control new seller in the set (table 4 column 1). We reach a similar conclusion when we incorporating the possibility that consumers might place an order in later days. Having some new sellers in the consideration sets significantly increases the likelihood for consumers to make purchases by 1.9% in the next three days (column 2 in table 4) and 1.8% in the following week (column 3). Higher purchase probability indicates that consumers benefit from better matching in searches given their search efforts. Higher intensity of interactions with new sellers therefore leads to market expansion.

Comparing control and treated new sellers who appear in consumers' consideration sets, we find that interacting with these two types of new sellers leads to a similar increase in consumers' purchase probability. We test the difference in the estimated coefficients β^c and β^t and find that the difference is small and insignificant except for log total spending (column 4 table 4). Since the training increases new sellers' likelihood to appear in the consideration sets and improves their customer service quality, there is both selection and treatment effect. Overall we do not find adverse selection because of the training since interacting with treated new sellers is associated with better matching. More importantly, consumers get better matching outcomes as they interact with both types of new sellers compared to only incumbents in the consideration sets. The gap in purchase probability implies that there could be significant frictions that hinder new sellers' growth.

Impacts on Search Interacting with new sellers could affect consumers' search and browsing behaviors on the platform. On the one hand, consumers could spend less effort in search when they encounter higher quality new sellers. As a result, consumers will visit fewer sellers in these search sessions. On the other side, a better experience could induce consumers to search more, both within the session and beyond. Since encountering new sellers is rare, consideration sets with these sellers are larger (figure A5). Without access to the actual page rank data, we test for the potential impacts on search intensity by relaxing the constraint in the baseline specification. In the baseline specification, we control for consumers' search efforts with α_s fixed effect. We use an alternative specification to include only search keyword-date fixed effects. This way, we compare variations within consumers and search keywords but varying search efforts. Table A13 presents the results. Without controlling for search intensity, the effect of encountering new sellers on boosting overall purchase probability is larger, at 3.9% for having treated new sellers and 4.4% for having control new sellers in the set (column 1).

4.4 Validation: Allocation and Post-Purchase Experience

In this section, we provide further evidence validating that the previous results on interacting with new sellers lead to more purchases and market expansion are indeed driving

by new sellers, rather than changes in compositions of the incumbents. We also show that consumers' post-purchase experience are comparable whether they purchase from new sellers or incumbents.

Allocation Within Consideration Set To confirm that the higher purchase probability for consideration sets that include new sellers are indeed causing by new sellers, we restrict the consumer-seller pair sample to consideration sets where consumers make purchase from some sellers in the set. We then use the following specification:

$$Y_{ijs} = \beta^t \mathbf{T}_j + \beta^c \mathbf{C}_j + \mathbb{X}_j \gamma + \alpha_{is} + \epsilon_{ijs} \tag{4}$$

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 the seller j belongs to the 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, the number of products offered and its ratings.

The estimated results are summarized in table 5. By restricting the sample 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 in the treatment group, over incumbents that appear in their consideration sets. Specifically, on the day of the visit, consumers are 5.9% more likely to choose a treated new seller than an incumbent in the same consideration set (column 2). Comparing treated new sellers with control new sellers, consumers' purchase probability with the former is significantly higher by about 4.2% (column 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. 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 pair sample, we show that training benefits the consumers by improving their matching quality and confirm that the improvement is due to consumers interacting with new sellers. The results shut down the potential negative selection induced by the training. As discussed in section **??**, training helps treated new sellers to attract more sellers to their sites. Therefore, from the consumers' perspective, more consumers visit treated new sellers and treated new sellers are more likely to appear in consumers' consideration sets. The business stealing effect implies that control new sell-

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 seller where consumers made a purchase 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 the specific 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.

ers and incumbents might be less likely to appear in the consideration sets. Hence, by changing the compositions of sellers that consumers visit, the training could cause market expansion and market reallocation. Such a shift in the allocation of consumers' visit likelihood could benefit the consumers but may come at the costs of control new sellers and incumbents. Our evidence shows that new sellers have higher quality than incumbents so that reallocation improves overall efficiency. We quantify and decompose such impacts in section 5.

Post-purchase Experience In the next validation exercise, we assess the impacts on the post-purchase experience when consumers reallocate their purchases from incumbents to new sellers.

We again use the consumer-seller pair sample but include observations where consumers made purchases from the sellers on the same day of visit[51]. Each observation is a consumer-seller pair that belongs to a specific consideration set. Therefore, this sample contains the exact same consideration sets as the sample used in **??**. The final sample consists of 54,910 consumer-seller pairs. To characterize the post-purchase experience, we use indicators for whether or not consumers request for returns or refunds for the order as well as whether or not consumers make repeat purchase from the same seller in the subsequent month following the purchase. For 9.1% of purchases, consumers request refunds or returns. Consumers place repeat orders in the following month in 4.9% of the cases.

To evaluate the impacts of purchasing from new sellers on the experience, we use the following specification:

$$Y_{ijs} = \beta^t \mathbf{T}_j + \beta^c \mathbf{C}_j + \mathbb{X}_j\gamma + p_{ijs}\delta + \alpha_i + \alpha_s + \epsilon_{ijs} \tag{5}$$

Because consumers only purchase from one seller in the consideration set, we are unable to use consideration set fixed effect. Therefore, we control for consumer and search specific variations with fixed effect α_i and α_s instead. As in the previous section, the outcomes of interests are $\mathbf{T}_j$ and $\mathbf{C}_j$, indicators for whether or not the seller j belongs to the treatment or control group. Besides seller level controls $\mathbb{X}_j$ on seller j's pricing, the number of products offered and its ratings, we also include the size of the order p_{ijs} in the control.[52]

Table 6 shows how post-purchase experience may differ when consumers purchase from new sellers instead of the incumbents with specification 5. Overall, placing an order with new sellers does not significantly negatively impact consumers' likelihood to request returns or refunds (columns 1 and 2). Moreover, there is no difference in the likelihood of making repeated purchases from new sellers versus incumbents (column 3). These

[51] We check the results by expanding the sample to include the observations where consumers purchase from the sellers within a week of visit and the results are consistent.

[52] We could not add consumer level controls because there are very limited variations left after controlling for consumer fixed effect. Hence the control variables only include sellers' characteristics.

Table 6: Effect on Quality of Purchase

	Dependent variable:		
	Refund	Return	Repeat Purchase
	(1)	(2)	(3)
Treated Seller	0.0004	0.0003	0.006
	(0.005)	(0.003)	(0.005)
Control Seller	-0.009**	-0.002	0.003
	(0.004)	(0.003)	(0.005)
Incumbent Mean	0.072	0.025	0.051
Treatment - Control	0.0099*	0.0019	0.0028
	(0.0055)	(0.0035)	(0.0062)
Observations	54,910	54,910	54,910
R^2	0.844	0.821	0.730

Notes: Sample restricted to set of consumers appearing in at least 2 sets, sets with at least one new seller and consumer-seller pairs where consumers actually placed orders. Control variables include consumer's recent spending and searching, consumers' ratings (a proxy for their cumulative experience on the platform) 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.

results show that while consumers are more likely to purchase from new sellers, their post-purchase experience with the new sellers is no worse than that with incumbents. The difference between interacting with treated and control new sellers on the purchase quality is small. Hence, higher purchase probability does not come at the cost of lowering purchase quality and there is no evidence for negative selection.

5 Decomposition of the Impacts of the Training

Motivated by the 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 uses variations in consumers' choices to identify the primary sources of frictions: the mismatch between sellers' quality and the number of visitors they manage to attract. The model uses a flexible function to characterize the platform's matching rule based on sellers' performance, i.e. lagged number of visitors and the conversion

rate[53].

To evaluate the welfare implications of the training, we consider how the training changes treated sellers' likelihood to appear in consumers' consideration sets that affects matching quality and welfare. Because of the training, treated new sellers invest in marketing to attract more visitors to their sites, leading to business stealing from control new sellers and incumbents. That is, training changes the composition of sellers in consumers' consideration sets. The spillover on the non-treated new sellers is limited because most of the active sellers on the platform are incumbents[54]. In the model, we take consumers' search and browsing behaviors as given and only consider the market reallocation among different types of sellers[55].

5.1 Model Setup

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

$$\max_{j \in K_i} U_{ij} = V_{ij} + \epsilon_{ij} = x_j\beta - \alpha p_j + \xi_j + \epsilon_{ij} \tag{6}$$

p_j is the 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. A seller with higher quality ξ_j yields higher utility for all consumers visiting her site. The main source of friction the model cap-

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

[54] If reallocation occurs randomly among all the sellers, then by chance, most of the market reallocation will come from the incumbents.

[55] On the one side, treated new sellers are more likely to appear in consumers' search sets, resulting in market reallocation from incumbents and control new sellers to treated new sellers. By closing the knowledge gap, training helps to ensure that new sellers can participate in the competition for consumers' attention. On the other hand, consumers may change their search behaviors in response to the changing composition of the sellers they visit. Consumers will have lower search costs if the quality of (initial) matches improves, 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 search more in the future, raising the likelihood of purchases from other sellers. This channel could potentially benefit all sellers if the market. Comparing to the impacts on market reallocation and changes in the consideration sets on consumers' purchases, changes in consumers' search behaviors are second-order.

tures comes from the mismatch between sellers' quality ξ_j and their likelihood to appear in consumers' consideration sets. Since we use post-treatment data to estimate ξ_j, it captures both sellers' underlying quality and effect of the treatment. We do not explicitly model treatment effect. Based on the reduced form results, treated new sellers could have higher estimated ξ_j because of their improved customer service. ϵ_{ij} is the i.i.d. consumer-seller idiosyncratic preferences that reflect the unobserved components affecting consumers' decisions.

We assume that ϵ_{ij} follows a type I extreme value distribution and consumers' outside option of not purchasing from any seller have zero utility. We get the following familiar logit formulation

$$P_{ij} = \frac{\exp(x_j\beta - \alpha p_j + \xi_j)}{1 + \sum_{k \in K_i} \exp(x_k\beta - \alpha p_k + \xi_k)} \tag{7}$$

where P_{ij} is consumer i's 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 the number of products offered in x_j could correlate with ϵ_{ij}, which would bias the estimates. To address this concern, we use a set of instruments to jointly determine the number of products posted prod_{jt} and pricing level p_{jt} with

$$\begin{bmatrix} p_{jt} \\ \text{prod}_{jt} \end{bmatrix} = \mathbf{Z}\beta^{fs} + \xi_j^{fs} + \xi_t^{fs} + \xi_s^{fs} + \epsilon_{jt} \tag{8}$$

ξ_j^{fs}, ξ_t^{fs} and ξ_s^{fs} are seller, date and sector fixed effect. The instruments Z capture the stringency of the platform's rule enforcement. These instruments include frequencies of rule violations and shares of sellers identified as frequent rule violators in the corresponding sectors[56]. The most common rule violations include infringing on intellectual property rights, selling counterfeits and providing false or misleading product information. The platform enforces comprehensive rules to ensure implementation of relevant government regulations and to maintain a well-functioning market. In the cases of rule violations, the platform might 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 rules in responses to changes in the business environment. We exploit changes in the stringency of rule enforcement at the sector-month level. When the platform strengthens the rule enforcement, sellers could be more cautious about posting products, charging ex-

[56]We use the number of visitors each seller attracts as the weights when we aggregate these instruments to sector level. Rule violations come in different degrees of severity which we separately calculate.

treme prices or engaging in unruly promotions. On the other side, the rule-obeying sellers are better positioned when the platform imposes stricter regulations on their unscrupulous competitors.

Matching The algorithm that matches sellers and consumers is the most important device that the platform has to improve consumers' experience and to support the promising sellers. We simplify the complex matching rules used in the search and recommendation algorithm by focusing on the reliance on the previous period's conversion rates and the number of visitors. In the matching, conversion rate directly reflects seller-specific consumer demands affected by sellers' underlying quality ξ_j. The 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 and training participation do not directly enter the matching function, but they can affect the lagged results. We capture the evolution of the number of visitors over time with the following model:

$$\mathcal{T}_{jt} = f(\mathcal{T}_{jt-1}, C_{jt-1}, C_{jt-2}) \tag{9}$$

$\mathcal{T}_{jt}$ is the current period total number of visitors for the seller j and $\mathcal{T}_{jt-1}$ is the previous period's number of visitors[57]. C_{jt-1} and C_{jt-2} are the 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 baseline 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, where we explore seller level variations. x_j includes the number of products sellers offer and sellers' ratings. p_j is the average price level that sellers charge[58]. To account for sector-level heterogeneity, we add a sector-specific intercept ξ_s in the baseline model and enrich the baseline model by estimating the 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 the average frequency of rule violations and the share of sellers labeled as frequent rule violators in sector s over the previous 30-

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

[58] 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 prices reflect the relative popularity of products sellers offer as well as any sales or promotions sellers offer.

day period. In the baseline model, we do not explicitly account for temporal variations. More detailed estimation procedures are described in B.2.

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

$$T_{jts} = f(T_{j,t-1}, C_{j,t-k}) + g(I_{j,t-k}(t-k=0)) + h(I^c_{j,t-k}(C_{j,t-k}=0)) + \xi^{trf}_k + \xi^{trf}_t + e_{jts} \quad (10)$$

$C_{j,t-k}$ are lagged conversion rates in past periods $k = 1, 2$, $I_{jt-k}(t-k=0)$ is an indicator for the initial two periods[59] 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 5 plots the distribution of the estimated sector fixed effect ξ_s. Panel B of figure 5 plots the distribution of the price elasticity for sellers in the sample. The average price elasticity is -0.22.[60]

The main parameters of interest are ξ_j for seller $j \in J$. These seller fixed effects can be interpreted as a quality measure for their products and services. Figure 5 plots the distribution of ξ_j across all sellers. Focusing on those with at least one purchase records (28% of all sellers), new sellers have significantly higher ξ_j than incumbents (Table 7). Among sellers with purchase records, the estimated quality of new sellers is 23.1% higher than incumbents. This means that the average product offered by new sellers provide 65-cent higher consumer gain. 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.

[59] Sellers have no past history on number of visitors and conversion rate in the initial periods hence are subject to different matching rules.

[60] The elasticity we estimate here is much smaller than the typical elasticity observed in the literature broda2006. 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. The average standard deviation of prices among all sellers that some consumers visited is 3.14 times higher than the average standard deviation of prices among sellers in the consumer-specific consideration sets.

Figure 5: Distribution of Estimated Parameters

(a) ξ_s (b) Elasticity

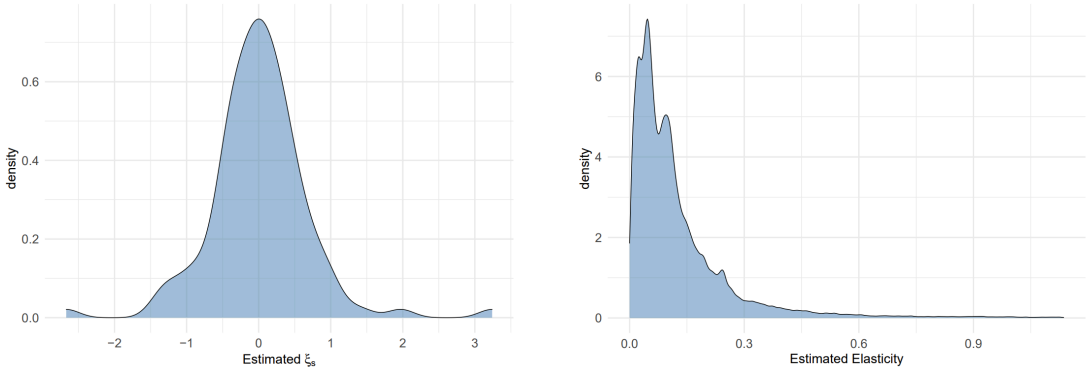


Table 7: Sellers' Characteristics and Estimated Type ξ

	Dependent variable:
	ξ_j
Control	0.142***
	(0.032)
Treatment	0.109***
	(0.055)
Constant	0.615***
	(0.007)
Observations	14,624

Notes: Distribution of estimated ξ_j on a sub-sample of new sellers and incumbents conditional on making at least one sale. See appendix B.1 for details.

5.4 Counterfactual: Welfare of the Training

We analyze the welfare consequences 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 any 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 consumer-search session pairs and sample sellers from the sellers' pool associated with the search keyword according to sampling weights. We match the number of sellers sampled with the observed number of sellers each sampled consumer visited in the corresponding search sessions. Therefore in the counterfactual, we only vary the 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 In the baseline version, sellers' sampling weights are 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 the 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 8 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[61].

Welfare Impact of Reducing GB with Training In this exercise, we use our structural model to compute the welfare impact of training, focusing on its ability to reduce the growth barrier. In particular, we calculate a counterfactual world in which trained sellers

[61] The sample we used in this estimate is not the same as the sample used in 4 because of the model specification. See appendix B.1 for more details on how the sample is constructed.

still faced the same growth barrier as their counterpart in the control group. This reduces their likelihood to appear in consideration sets.[62]

Table 8: Welfare and Market Share Decomposition

Revenue	**CS**	**Market Share**		
		Treatment	Control	Incumbent
-0.05	-0.07	-7.71	0.17	0.13
(0.05)	(0.03)	(0.63)	(0.11)	(0.02)

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. Bootstrapped standard errors are in parentheses. Revenues and consumer surpluses are changes from baseline level revenues and consumer surpluses respectively.

Overall, only 4.3% of the sellers are new, so that trained new sellers make up only 0.25% of all sellers. However, removing the training program from such a small portion of the market is enough to reduce total revenue reduces by 0.05%, with consumer welfare dropping by 0.07%. In particular, treated seller suffer 7.7% drop in market share, obtaining a smaller share of a smaller pie.[63] Both control-group and incumbent sellers see moderate increase in market share.

6 Conclusions

In this chapter, 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 search and browsing data, we show that the business training helps new sellers gain more exposure 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 purchases' quality. Using a structural model where we highlight the mismatch between the number of visitors that sellers attract and their quality, we show

[62]We randomly remove a subset of training participants from the sampling pool that consumers could choose from, until the likelihood of treated new sellers' appearance in the consideration sets matches those of control new sellers. See Appendix B.3 for more detail.

[63]Percentage point are hidden at data provider's request.

that the training increases consumers' welfare and total revenues by limiting the extent of misallocation. The improved matching quality and new sellers' higher service quality could lower consumers' search costs as well, which we do not account for in the current analysis. Our results also motivate the platform's strategic choices to offer interventions as such to support its participants. 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' engagement with online marketing also contributes to the platform's profits.

Our findings provide one of the first sets of direct empirical evidence on the consumer welfare implications of an intervention that supports the subset of firms. Consumers do not experience adverse selection in this context because the training reduces market frictions and improves 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. Doing so would align with platforms' incentives as profit-maximizing firms in both the short and long run. Although large e-commerce platforms' looming market power should not be overlooked, these platforms indeed create enormous opportunities for SMEs. They have incentives and the capacity to take more active roles to foster the efficiency and equality in online marketplaces they host.

References

Agarwal, Nikhil, Susan Athey, and David Yang (2009). "Skewed bidding in pay-per-action auctions for online advertising". In: *American Economic Review* 99.2, pp. 441–47.

Anderson, Stephen J., Rajesh Chandy, and Bilal Zia (Dec. 2018). "Pathways to Profits: The Impact of Marketing vs. Finance Skills on Business Performance". In: *Management Science* 64.12, pp. 5559–5583.

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

Atkin, David, Amit K Khandelwal, and Adam Osman (2017). "Exporting and firm performance: Evidence from a randomized experiment". In: *The Quarterly Journal of Economics* 132.2, pp. 551–615.

Backus, Matthew, Thomas Blake, Brad Larsen, and Steven Tadelis (2020). "Sequential bargaining in the field: Evidence from millions of online bargaining interactions". In: *The Quarterly Journal of Economics* 135.3, pp. 1319–1361.

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

Barach, Moshe A, Joseph M Golden, and John J Horton (2020). "Steering in online markets: the role of platform incentives and credibility". In: *Management Science* 66.9, pp. 4047–4070.

Bärtl, Mathias (2018). "YouTube channels, uploads and views: A statistical analysis of the past 10 years". In: *Convergence* 24.1, pp. 16–32.

Berman, Nicolas, Vincent Rebeyrol, and Vincent Vicard (2019). "Demand learning and firm dynamics: evidence from exporters". In: *Review of Economics and Statistics* 101.1, pp. 91–106.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts (2013). "Does management matter? Evidence from India". In: *The Quarterly Journal of Economics* 128.1, pp. 1–51.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar (2018). "The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in Mexico". In: *Journal of Political Economy* 126.2, pp. 635–687.

Calderon, Gabriela, Jesse M Cunha, and Giacomo De Giorgi (2020). "Business Literacy and Development: Evidence from a Randomized Controlled Trial in Rural Mexico". In: *Economic Development and Cultural Change* 68.2, pp. 507–540.

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

Cullen, Zoë and Chiara Farronato (2020). "Outsourcing tasks online: Matching supply and demand on peer-to-peer internet platforms". In: *Management Science*.

Cusumano, Michael (2010). "Technology strategy and management The evolution of platform thinking". In: *Communications of the ACM* 53.1, pp. 32–34.

Dinerstein, Michael, Liran Einav, Jonathan Levin, and Neel Sundaresan (2018). "Consumer price search and platform design in internet commerce". In: *American Economic Review* 108.7, pp. 1820–59.

Eisenmann, Thomas, Geoffrey Parker, and Marshall W Van Alstyne (2006). "Strategies for two-sided markets". In: *Harvard business review* 84.10, p. 92.

Farronato, Chiara, Andrey Fradkin, Bradley Larsen, and Erik Brynjolfsson (2020). *Consumer protection in an online world: An analysis of occupational licensing*. Tech. rep. National Bureau of Economic Research.

Fitzgerald, Doireann, Stefanie Haller, and Yaniv Yedid-Levi (2016). *How exporters grow*. Tech. rep. National Bureau of Economic Research.

Foster, Lucia, John Haltiwanger, and Chad Syverson (2016). "The slow growth of new plants: Learning about demand?" In: *Economica* 83.329, pp. 91–129.

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

Ghose, Anindya and Sha Yang (2009). "An empirical analysis of search engine advertising: Sponsored search in electronic markets". In: *Management science* 55.10, pp. 1605–1622.

Gourio, Francois and Leena Rudanko (2014). "Customer capital". In: *Review of Economic Studies* 81.3, pp. 1102–1136.

Hagiu, Andrei (2014). "Strategic Decisions for Multisided Platforms". In: *MIT Sloan Management Review* 55.2, p. 71.

Hagiu, Andrei and Julian Wright (2020). "Platforms and the exploration of new products". In: *Management Science* 66.4, pp. 1527–1543.

Horton, John J (2017). "The effects of algorithmic labor market recommendations: Evidence from a field experiment". In: *Journal of Labor Economics* 35.2, pp. 345–385.

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

Jin, Yizhou and Zhengyun Sun (2019). "Information Acquisition and the Return to Data". In: *Harvard University Working Paper*.

Kim, Hyunjin (2019). "The Value of Competitor Information: Evidence from a Field Experiment". In: *Harvard Business School Working Paper*.

Lewis, Gregory and Georgios Zervas (2019). "The supply and demand effects of review platforms". In: *Available at SSRN 3468278*.

McKenzie, David (forthcoming). "Small business training to improve management practices in developing countries: Reassessing the evidence for "training doesn't work"". In: *Oxford Review of Economic Policy*.

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

McKenzie, David and Christopher Woodruff (2014). "What are we learning from business training and entrepreneurship evaluations around the developing world?" In: *The World Bank Research Observer* 29.1, pp. 48–82.

Milgrom, Paul R and Steven Tadelis (2018). "How Artificial Intelligence and Machine Learning Can Impact Market Design". In: *NBER Working Paper* w24282.

Moshary, Sarah (2021). "Advertising Effects in Equilibrium". In:

Muralidharan, Karthik and Paul Niehaus (2017). "Experimentation at scale". In: *Journal of Economic Perspectives* 31.4, pp. 103–24.

Parker, Geoffrey G, Marshall W Van Alstyne, and Sangeet Paul Choudary (2016). *Platform revolution: How networked markets are transforming the economy and how to make them work for you*. WW Norton & Company.

Piveteau, Paul (2016). *An empirical dynamic model of trade with consumer accumulation*. Tech. rep. Working Paper.

Reshef, Oren (2019). *Smaller Slices of a Growing Pie: The Effects of Entry in Platform Markets*. Tech. rep.

Sahni, Navdeep S and Harikesh S Nair (2020). "Does Advertising Serve as a Signal? Evidence from a Field Experiment in Mobile Search". In: *The Review of Economic Studies* 87.3, pp. 1529–1564.

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

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

Ursu, Raluca M (2018). "The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions". In: *Marketing Science* 37.4, pp. 530–552.

Zhang, Dennis J, Hengchen Dai, Lingxiu Dong, Qian Wu, Lifan Guo, and Xiaofei Liu (2019). "The value of pop-up stores on retailing platforms: Evidence from a field experiment with alibaba". In: *Management Science* 65.11, pp. 5142–5151.

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:

$$Y_{imcs} = \sum_{m=1}^{9} \beta^m Treatment_i M_{im} + \alpha_m + \alpha_c + \alpha_s + \epsilon_{imcs} \tag{11}$$

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 performance is relatively short-lived. Figure A3 presents the estimated coefficients $\{\beta^m\}_{m=1}^9$ in specification (11) 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 the 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 targets basic operations and marketing, we expect sellers with limited previous exposure to e-commerce to benefit more from the training as it helps to close their 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[64].

[64]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

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

$$Y_{imcs} = \beta Treatment_i + \gamma Type_i + \delta Treat_i \times Type_i + \alpha_m + \alpha_c + \alpha_s + \epsilon_{imcs} \quad (12)$$

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 A9 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. The slightly surprising result is that there is no differential treatment effect for sellers with different preparedness levels. Compared to the rest, sellers who post products on the first day of entry could be better prepared or more experienced. Hence, these sellers might find the basic part of the training less useful, yet we do not find such results[65].

To evaluate the impacts of the offline business environment, we group the sellers based on their registered locations[66]. Table A10 presents the results on heterogeneous treatment effect by sellers' locations on log monthly revenues. γ captures sellers' average performance in different parts of the country compared 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 southern coastal provinces significantly outperform the rest, while those coming from the less-developed western part lagged. The 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 (see table A1)[67]. Therefore, even the training program offers the 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.

convert.

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

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

[67] Training participants are significantly less likely to be sellers from the western provinces.

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 **??**. 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 **??** is that we also include consideration sets that contain incumbents only, as opposed to just new sellers. Because of the computational constraint and setup of the model, we cannot 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 control) from the pool of new sellers that appear in the full consumer-seller pair sample.

2. For each new seller, obtain all the incumbents 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 and 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 they are more likely to appear in multiple consideration sets. We use traffic to refer to the number of visitors a seller obtains in 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 the seller level, even though in the search process, consumers access specific sellers' sites by searching specific products rather than the 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 the ones that direct them to sellers' sites. Since we do not observe consumers' browsing history in the stores, we cannot fully capture such a process. Instead, since we are predominantly interested in the

seller level characteristics, we aggregate all the purchase and browsing behaviors to the consumer-seller level, rather than to the consumer-seller-product level. The number of products offered and the ratings are measured at the seller level. Pricing is the quantity weighted prices of all products that the sellers offer.

The instruments we used for sellers' pricing and the number of products offered are variables that capture the platform's rule enforcement intensity. 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 7 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 8. 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 3 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 are 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 a cross validation method and calculate the average residuals on the training sample. Table A14 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 those 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 $w_{js}^0 = \frac{\hat{T}_j}{\sum_{k \in S} \hat{T}_k}$. We use w_{js}^0 as the baseline sampling weights. With the estimated sampling weights for all the sellers in the pool, we randomly sample sellers from a consumer-set specific pool where the number of sellers is the same as the number of sellers the consumer visits during that particular search session. Therefore, the only part that is changed in the counterfactual analysis is the composition of sellers in consumers' consideration sets. In contrast, the 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. To quantify the impacts of the training, we assume 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 a 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 training's influence, 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 other sellers' sampling weights 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 and different types of sellers' market shares if they appear in the consideration sets with some new sellers under each counterfactual scenario. We then calculate the changes in market shares from the baseline market allocation.

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 we collected 566 responses. Detailed results are presented in table A16. 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 anecdotal descriptions of typical sellers, these sellers appear to have higher than average ownership of manufacturing factories (32.5%) and offline stores (19.2%).

The survey shows that even among the training participants, sellers differ in terms of their background, experience, education and financial resources. However, while the 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 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, and 68.2% 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[68].

[68] 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: 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 the 1st Day	0.213	0.212	0.0005	0.245	0.204	0.04***
	(0.409)	(0.409)	(0.44)	(0.430)	(0.403)	(16.04)
# 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)
# Visitors	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 shows 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. Columns 6 shows the comparison between training participants and non-participants. For the second row (female owner) the value in parentheses shows the estimate with a further sample restriction that includes only those 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 A2: Treatment Effects on Monthly Revenues

	Dependent variable:						
	Monthly Revenues						
	Indicator	Raw	Log	I.H.S.	Winsorized		
					99th	99.5th	99.9th
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	0.002**	-939.245	0.017***	0.018***	25.436**	26.175	85.084
	(0.001)	(1,447.148)	(0.006)	(0.007)	(12.478)	(22.308)	(54.321)
Dep Var Mean	0.19	7018.16	1.39	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	6,409,062
R^2	0.152	0.0001	0.132	0.134	0.027	0.024	0.019
Adjusted R^2	0.152	0.00004	0.132	0.134	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. In column 4, the revenue is transformed with inverse hyperbolic sine. We calculate thresholds for winsorization by the relative month since registration. 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: 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 posted the first product. Columns 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 A4: Main Results Among Sellers Ever Posted Products

	Dependent variable:					
	Post Products	Any Revenues	Revenues	# Visitors	# Buyers	Conversion Rate
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.001	0.002*	0.024***	0.018**	0.012**	0.0001
	(0.001)	(0.001)	(0.010)	(0.008)	(0.005)	(0.0002)
Dep Var Mean	0.59	0.32	2.32	2.82	0.95	0.04
Observations	6,409,062	3,802,509	3,802,509	3,802,509	3,802,509	2,593,762
R^2	0.380	0.100	0.081	0.118	0.064	0.045
Adjusted R^2	0.380	0.100	0.081	0.118	0.064	0.045

Notes: The outcome variable for column 1 is an indicator for seller ever posting a product using the full sample. In column 2 to 5 the sample is restricted to sellers that ever post any products in 9-month. In column 6 the sample is further restricted to sellers who have visitors during the month. 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: IV Results on Main Outcomes: Any Tasks

	Dependent variable:					
	Take-up Training	Any Revenues	Revenues	# Visitors	# Buyers	Conversion Rate
	(1)	(2)	(3)	(4)	(5)	
Treatment	0.257***					
	(0.001)					
Take-up Training		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	0.153	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 is 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 as the instrumental variable. All specifications are two-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 A6: IV Results on Main Outcomes: Number of Tasks Taken-up

	Dependent variable:				
	Any Revenues	Revenues	# Visitors	# Buyers	Conversion Rate
	(1)	(2)	(3)	(4)	(5)
# Tasks	0.028***	0.284***	0.208***	0.139***	0.002
	(0.011)	(0.088)	(0.077)	(0.043)	(0.002)
Dep Var Mean	0.19	1.39	1.73	0.57	0.04
Observations	6,409,062	6,409,062	6,409,062	6,409,062	2,593,762
R^2	0.155	0.136	0.209	0.109	0.049

Notes: The first stage specification is the following:

$$Tasks_{it} = \sum_{k=1}^{9} \gamma_k Treatment_i \mathbf{I}_{t=k} + \alpha_t + \alpha_s + \alpha_c + \epsilon_{it}$$

where $Tasks_{it}$ is number of tasks taken-up for seller i during month i. $\mathbf{I}_{t=k}$ is an indicator for being in the kth month since registration. All specifications are two-stage least square results using treatment assignment interacting with dummies for relative month as the instrument. 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 A7: 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 an indicator of take up of at least during the sample period. 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 A8: 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 calculates and assignes 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 the frequency that sellers violate the platform's rules, see more details in appendix B.1, descriptions of the instruments. Share of good reviews is the share of good reviews out of all the reviews that sellers get. Vast majority of the reviews are positive. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table A9: 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 A10: 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 A11: Summary Statistics: Consumer-Search Session Sample

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Number of Sellers Browsed Per Session	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 A12: 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 A13: Visiting New Sellers and Consumers' Purchase without Search Effort Control

	Dependent variable: Purchase				
	Purchase			Log Spending	Log Order Size
	Same Day	In 3 Days	In a Week	Same Day	
	(1)	(2)	(3)	(4)	(5)
Any Treated Seller	0.007***	0.010***	0.011***	0.021***	0.031***
	(0.001)	(0.002)	(0.002)	(0.007)	(0.007)
Any Control Seller	0.008***	0.009***	0.0011***	0.033***	0.042***
	(0.001)	(0.001)	(0.001)	(0.006)	(0.006)
Incumbent Mean	0.18	0.21	0.22	0.76	0.78
Treatment - Control	-0.0049	0.00061	0.00019	-0.013	-0.011
	(0.0017)	(0.0019)	(0.0019)	(0.0079)	(0.0084)
Observations	1,381,273	1,381,273	1,381,273	1,381,273	1,381,273
R^2	0.662	0.649	0.637	0.683	0.674

Notes: All regressions include search keyword-date 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 A14: 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 the out-of-sample data.

Table A15: 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 ordered in sequence of priorities. Each task is triggered by a particular indicator. For the outcome based tasks, comparisons 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 A16: 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 wholesale markets	21.7%
Some College	28.0%	Distribution/brand subsidiary	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: Training Participation: Tasks Take-up and Completion

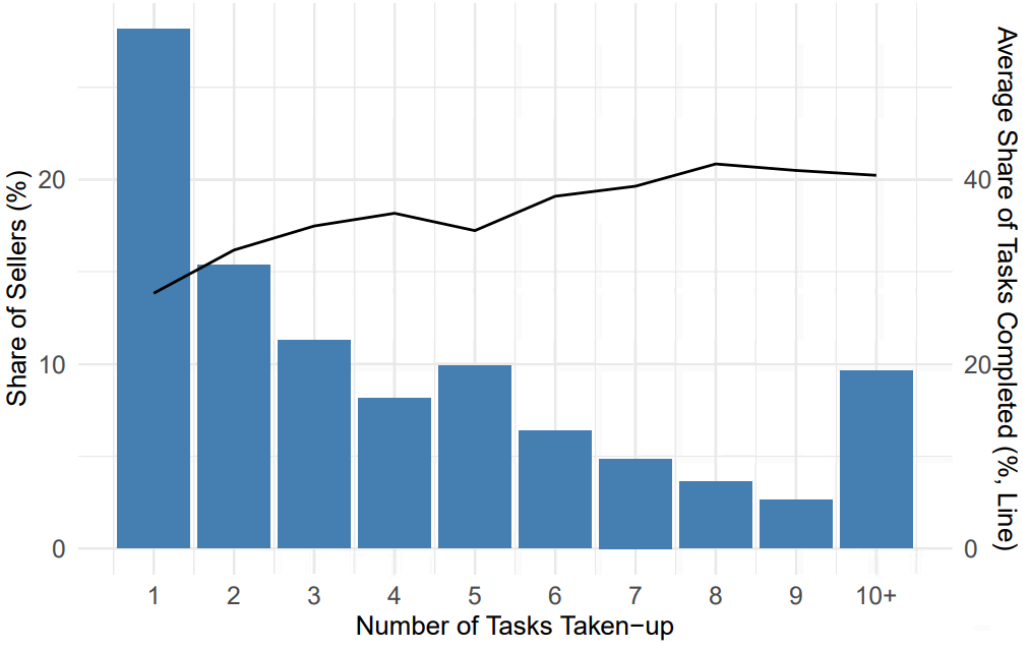


Figure A2: Quantile Treatment Effect on Revenue Over Time

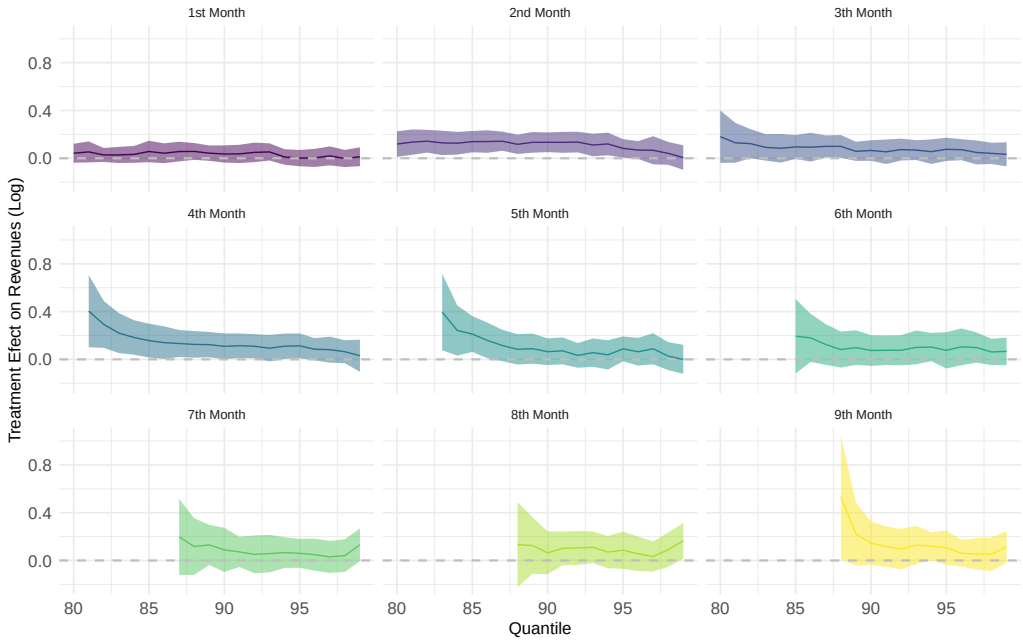
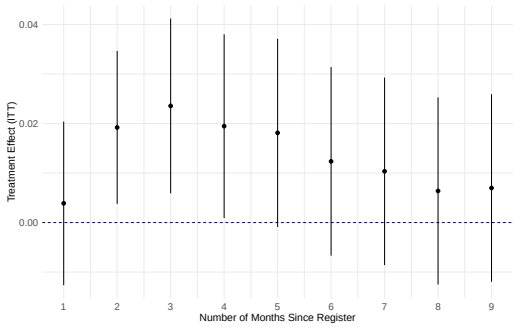
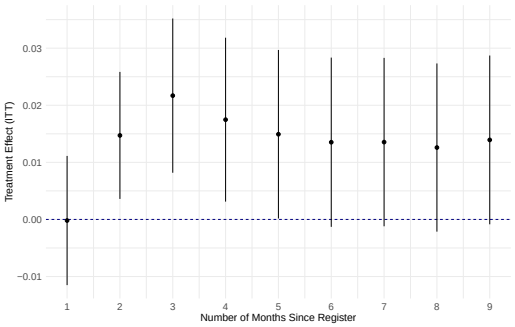


Figure A3: Long Term Treatment Effect on Main Outcomes

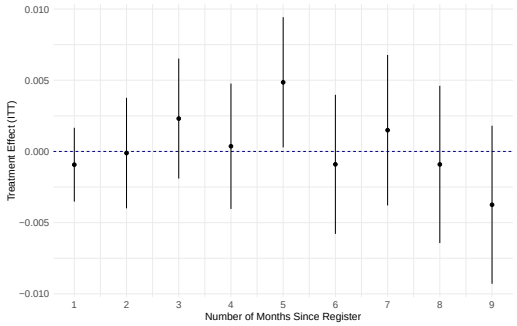
(a) Revenues



(b) Number of Visitors



(c) Conversion Rate



(d) Number of Buyers

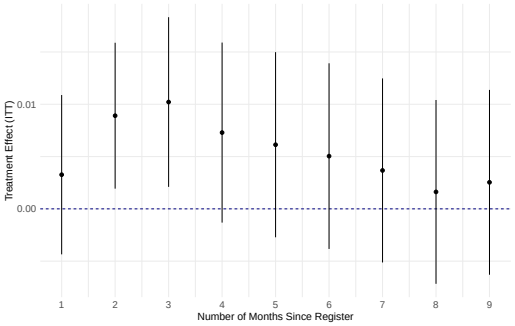


Figure A4: Procedures of Constructing the Consumer Sample

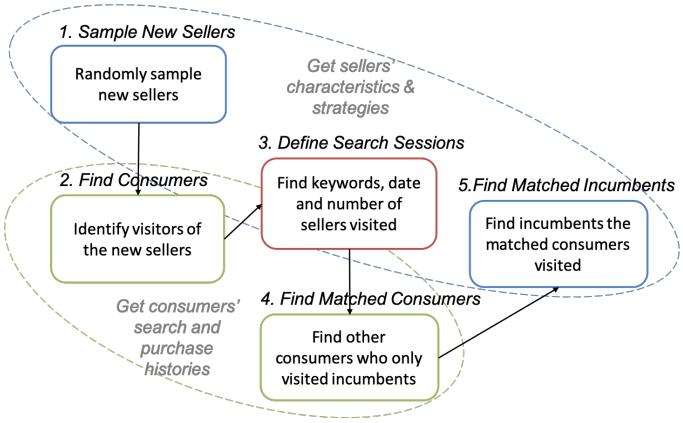


Figure A5: Distribution of Size of Search Sets

