NBER WORKING PAPER SERIES

SEARCH AND INFORMATION FRICTIONS ON GLOBAL E-COMMERCE PLATFORMS: EVIDENCE FROM ALIEXPRESS

Jie Bai
Maggie Chen
Jin Liu
Daniel Yi Xu

Working Paper 28100
http://www.nber.org/papers/w28100

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2020

We thank Chang-tai Hsieh, Costas Akolakis, David Atkin, Lauren Bergquist, Ben Faber, Gordon Hanson, Asim Khwaja, Pete Klenow, Meredith Startz, Tianshu Sun, and seminar and conference participants at the Berkeley Economics, BREAD/CEPR/STICERD/TCD, Uni-versity of Chicago, Conference, Harvard Kennedy School, IPA SME Working Group Meeting, Michigan Economics, and USC for helpful comments. We thank Chengdai Huang, Haoran Zhang, and Qiang Zheng for excellent research assistance. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Jie Bai, Maggie Chen, Jin Liu, and Daniel Yi Xu. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Search and Information Frictions on Global E-Commerce Platforms: Evidence from AliExpress
Jie Bai, Maggie Chen, Jin Liu, and Daniel Yi Xu
NBER Working Paper No. 28100
November 2020
JEL No. F14,L15,O12

ABSTRACT

Global e-commerce platforms present new export opportunities for small and medium-sized enterprises in developing countries by significantly lowering the entry barriers of exporting. However, the lack of market selection can lead to a large number of online firms competing for consumers' attention, resulting in severe congestion in consumers' search process. When firms' intrinsic quality is not perfectly observed, these search frictions can further slow down the resolution of the information problem and hinder market allocation towards better firms. In this paper, we investigate how search and information frictions shape firm dynamics and market evolution in global e-commerce. Using detailed data from AliEpxress as well as a rich set of self-collected objective quality measures, we provide stylized facts that are consistent with the presence of search and information frictions. Moreover, using a randomized experiment that offers exogenous demand and information shocks to small prospective exporters, we establish that firms with larger past sales have an advantage in overcoming the search friction and generating future orders. This indicates that initial demand shocks could confound firms' true quality in determining firm growth and the long-run market structure. We construct and estimate an empirical model of the online market that are consistent with our descriptive and experimental findings and use the model to quantify the extent of demand-side frictions. Counterfactual analyses show that alleviating information frictions and reducing the number of firms can help to improve allocative efficiency and raise consumer welfare.

Jie Bai
Harvard Kennedy School
79 JFK Street
Cambridge, MA 02138
and NBER
jie_bai@hks.harvard.edu

Maggie Chen
Dept. of Economics
George Washington University
2115 G ST, NW, #367
Washington, DC 20052
xchen@gwu.edu

Jin Liu
Department of Economics
New York University
jl7767@nyu.edu

Daniel Yi Xu
Department of Economics
Duke University
213 Social Science Bldg
419 Chapel Drive
Box 90097
Durham, NC 27708-0097
and NBER
daniel.xu@duke.edu

1 Introduction

E-commerce sales have grown tremendously in recent years, reaching $2.9 trillion in 2018 and 12 percent of the total global retail sales (Lipsman, 2019). Within e-commerce, cross-border sales have grown two times faster than domestic sales, and nearly 40 percent of online buyers completed a cross-border transaction in 2016 (Pitney Bowes, 2016). By extending market access beyond geographical boundaries, global e-commerce platforms present a promising avenue for small and medium-sized enterprises (SMEs) in developing countries to enter into export markets. Furthermore, online exporting lowers many of the traditional barriers of offline exporting, including the needs of building export relationships and setting up distributional channels in destination countries.[1] Given these promises and the large market potential, numerous policy initiatives have been adopted worldwide to foster e-commerce growth (e.g, UNCTAD, 2016), with a specific policy target to onboard SMEs in developing countries to e-commerce platforms and allow them to tap into the global market.

Despite the rapid growth of global e-commerce, there is a lack of empirical evidence on the impact of such increased export opportunity on firm growth and market dynamics. While e-commerce platforms potentially expose prospective exporters to buyers around the world, the sheer number of firms operating on these platforms can create substantial congestion in consumer search. When firms' intrinsic quality is not perfectly observed, these search frictions can further slow down the resolution of the information problem and hinder market allocation towards better firms. In such an environment, initial demand shocks (or "luck")—as opposed to economic fundamentals such as firm productivity or product quality—can have a persistent impact on firms' long-run growth and market allocation. Understanding the role of these demand-side forces in firm growth and welfare can lead to important policy implications.

In this study, we investigate how search and information frictions shape firm dynamics and market evolution in global e-commerce. We first document descriptive evidence that is consistent with the existence of sizable search and information frictions in global online marketplaces. Next, we experimentally identify and demonstrate the role of accumulating demand in helping firms to overcome these frictions and generating future demand. Finally, we combine these elements to build and estimate a rich empirical model of the online export market. We use the model estimates to quantify the distinctive impacts of search and information frictions on firm growth, market allocation, and consumer welfare. Finally, we apply our model to shed light on policies that could facilitate the growth of promising export businesses beyond the initial onboarding stage and improve the overall market efficiency.

Our study is grounded in the context of AliExpress, a world-leading B2C cross-border e-commerce platform owned by Alibaba. We focus on the industry of children's t-shirts and collect comprehensive

[1] For example, AliExpress, one of the leading cross-border e-commerce platform that we study in this project, states on its website (`https://sell.aliexpress.com/__pc/4DYTFsSkV0.htm`): "Set up your e-commerce store in a flash, it's easy and free! Millions of shoppers are waiting to visit your store!"

data about sellers operating in this industry, including detailed seller-product-level characteristics and transaction-level sales records. We complement the platform data with a novel set of objective, multi-dimensional measures of quality, ranging from detailed product quality metrics to shipping and service quality indicators. These measures are collected by the research team based on actual online purchases and direct interactions with the sellers, as well as third-party assessments.

We begin by documenting a set of new stylized facts about the online exporters. First, we compare sales distribution within "identical-looking" product varieties. Interestingly, even after controlling for horizontal taste differences, meaningful dispersion in sales remains within identical variety groups, as opposed to "winner-takes-all". This finding is indicative of search frictions: buyers, upon arriving at the platform, face thousands of product offerings but can only sample a limited finite subset of all seller-listings. The relatively low fixed costs of operating in the online marketplace weakens the market selection mechanism and exacerbates congestion in consumer search, resulting in an excessive number of firms and product offerings in the online marketplace competing for consumers' attention. This raises the question of who gets to grow in the presence of the search problem. Next, we dive further into the potential determinants of growth and find that quality only weakly predicts sales. The "superstars", which we define as the largest seller in each product variety, do not necessarily have the highest quality (nor the lowest price). Intuitively, search friction introduces a random component in firm growth due to the consumer sampling process. Furthermore, when firms' intrinsic quality is not perfectly observed, such friction can further slow down the revelation of true quality, leading to potential market misallocation. Finally, we find robust evidence that current sales predict the speed of arrival for future sales. This implies that firms with larger past sales, hence higher visibility, have an advantage in overcoming the consumer's search friction and generating future orders. However, if information friction prevents a firm's visibility from being aligned with its fundamentals, such as quality or productivity, the same force could lead to allocative inefficiency. In particular, random demand shocks (or "luck"), as opposed to firm fundamentals, can have a persistent impact on firms' long-run growth. Over time, firm performance diverges; market allocation and consumer welfare depend crucially on the interactions of these demand-side forces.

Our interpretation of the stylized facts centers around a demand driven mechanism where each additional consumer order makes the selling firm more visible and hence helps the firm to overcome the search frictions faced by subsequent buyers. However, important unobserved supply-side actions (such as web positioning and advertising) could also exist and lead to similar reduced-form relationships between current sales and future sales. To further establish the empirical validity of the demand mechanism, we conduct an experiment in which we generate exogenous demand shocks to a set of small exporters via randomly-placed online purchase orders. The treatment allows us to isolate the impact of demand from unobserved supply-side confounding factors. Since how effectively the additional demand conveys the firm's true fundamentals depends critically on the severity of information friction, we further interact the order treatment with a review treatment about firms' product and shipping quality to examine the role of information provision. We track these firms over four months and collect high-frequency data on sales and prices as well as objective measures of quality. We find that the order treatment leads to a small but significantly positive impact on firms' subsequent sales. This demonstrates that indeed a key channel for firms to improve their visibility and grow in the online marketplace is by accumulating sales. Quantile analysis reveals, however, that the effect is mainly concentrated at the top: only a small fraction of sellers are able to take advantage of the initial demand shock and grow while the vast majority stay small. The size of the average treatment effect suggests that these demand-side frictions cannot be easily overcome by individual sellers' private efforts. In the meantime, we do not find any significant treatment effect from the reviews, suggesting that the online reputation mechanism may not function very effectively in the presence of large search friction. Intuitively, reviews only matter when a seller's listing is discovered by consumers, which is a rare event for small businesses due to their low visibility. Finally, we do not find significant heterogeneous treatment effect based on quality. This echoes the stylized fact that quality does not strongly predict growth in this market due to the search and information problems, which, combined, make it difficult for high-quality sellers to stand out.

All together, the descriptive and experimental findings are consistent with the presence of large search and information frictions and show that in such an environment demand shocks, as opposed to firm fundamentals, can affect the firm's future growth. Motivated by the reduced-form evidence, we next build a structural model of the online market incorporating these realistic frictions of the market. Our model features consumers' finite sample search to capture search frictions and the online review mechanism to capture information frictions. Our estimates imply that a consumer can sample only 0.2% of all seller-listings on the e-commerce platform.[2]. However, once a seller starts to make sales, the initial success in receiving orders substantially increases a seller's visibility. Compared with sellers who have made zero sales, striking a first order makes a seller 3.4 times more likely to end up in a subsequent consumer's search sample. On the other hand, the estimate of the review signal noise indicates substantial information frictions. The posterior uncertainty is only reduced by 7.5% after the first order. This implies that the reputation mechanism takes time to play its role: even if a seller gets sampled and successfully makes a sale, uncertainty regarding quality still remains and only resolves slowly. Combined, these findings highlight that search friction, interacted with information friction, can constitute an important hurdle for the growth of small prospective exporters. When we simulate a one-time demand treatment through the lens of the model, we find a smaller but quantitatively similar average treatment effect compared to what we obtain from the experiment.

We end with several counterfactual exercises to examine the distinctive roles of search and infor-

[2] We have close to $20,000$ seller-listings that sell children's t-shirt in our data sample. This implies a search sample size of 40 sellers

mation frictions in firm growth and market allocation and evaluate potential policy interventions using the estimated structural model. First, to shed light on information frictions, we remove the noise of the review signals. We find that doing so significantly shifts market share to high-quality sellers. The resulting consumer surplus is 12.7% higher compared to the baseline. Second, to examine the role of initial demand, we compare the baseline case in which initial demand is determined purely by luck versus a case where it is determined by quality. Remarkably, we find that just a ten-period difference in initial demand allocation generates a persistent long-run difference in market outcomes: the market share for sellers in the top quality quartile is 7.6 percentage points higher and consumer surplus is 7.4% higher. Finally, we investigate the impact of reducing search frictions by reducing the number of sellers operating on the platform. The results show that reducing the number of sellers can help mitigate the congestion in consumer search, thereby improving allocative efficiency and consumer welfare.[3] This result points out that just giving firms easy access to foreign markets alone may not be sufficient for generating sustained growth and can in fact exacerbate the search problem, resulting in market misallocation. Policies should be designed to help firms, especially new businesses, overcome the additional demand-side frictions. In the context of e-commerce, regulating entry, creating a premium market segment, and directing demand to promising newcomers could help facilitate growth and improve the overall market efficiency.

Our work contributes to several strands of the existing literature. Extensive work in international trade has studied the empirical patterns of new exporter dynamics in the offline setting. A common empirical pattern that emerges from micro data is that young exporting firms start small and have high turnover rates and those that survive experience rapid growth. Various theories have been proposed to explain these facts. They include firm learning (Arkolakis, Papageorgiou, and Timoshenko, 2018; Ruhl and Willis, 2017), demand accumulation(Foster, Haltiwanger, and Syverson, 2016; Piveteau, 2016; Fitzgerald, Haller, and Yedid-Levi, 2020), and seller search (Eaton et al., 2016). In contrast to these studies, our paper focuses on demand-side frictions and shocks, rather than the exporter's own decision, as the key driving force of firm and market dynamics in the online setting. In particular, unlike the offline export market, the fixed costs of operating are substantially lower in online marketplaces, significantly weakening the role of market selection. Our work shows how the lack of selection reduces consumer search efficiency and endogenously slows down the growth of high-quality exporters. Methodologically, we bring in new sources of variations to first experimentally identify the mechanisms underlying new

[3]Given the extensive numbers of varieties usually available in the online marketplace, our model and counterfactual analysis abstract from consumer welfare gains from additional varieties and focus instead on the implications of excessive entry for search friction. The potential adverse effect of excessive choices on individuals' decisions and utilities, albeit less explored in the economics literature, has been documented in studies of social psychology, often termed as "choice overload." Evidence from field and laboratory experiments (Scheibehenne, Greifeneder, and Todd, 2010; Chernev and Hamilton, 2009; McShane and Böckenholt, 2018) suggests that having too many options to choose from could lead to decision paralysis and negative psychological and behavioral effects. The literature emphasizes the importance of building a better choice architecture including helping individuals structure their search and assisting them in streamlining choices.

exporter's demand accumulation process and then formally model the realistic frictions of the market. Our findings also connect to the existing literature in trade that examines the roles of search and information frictions on market demand and seller reliability in explaining price variations and trade patterns (Allen, 2014; Macchiavello and Morjaria, 2015; Steinwender, 2018; Startz, 2018).[4]

Another complementary literature explaining exporter performance highlights the role of quality (see Verhoogen, 2020, for an excellent review). Since product or service quality is rarely observed in standard firm surveys, most of the earlier literature has focused on indirect measures of quality estimated based on market shares and prices, for instance, (Verhoogen, 2008; Khandelwal, 2010). We build on a growing body of development research that collects detailed information on quality for specific industries (e.g, Bai, 2016; Atkin, Khandelwal, and Osman, 2017; Hansman et al., 2020). Similar to these earlier works in offline settings, we document large variations in firm-product quality online. However, we find quality plays a less pronounced role in explaining exporter growth and long-run market shares. Our paper explains the disintegration of the customer accumulation process and firm fundamentals, such as quality, and underscores the potential sources of market share misallocation in the e-commerce context.

Last but not least, findings from our study also speak broadly to the development literature on interventions to help micro, small, and medium enterprises. Echoing the literature on productivity differences across firms, most of the earlier work has emphasized supply-side interventions, including providing credit access, quality inputs, and managerial training (e.g, De Mel, McKenzie, and Woodruff, 2008; Kugler and Verhoogen, 2012; Banerjee, 2013; Bloom et al., 2013). More recently, a growing set of work has begun to look at demand-side interventions. A closely related study to ours is Atkin, Khandelwal, and Osman (2017), which also studies the impact of foreign demand shocks on exporters, showing that firms respond to these demand shocks by improving quality through learning by doing. Rather than focusing on firms' own actions, we explore the impact of foreign demand shocks on search and information about the firm.

The remainder of the paper is organized as follows. Section 2 describes the empirical setting and data. Section 3 presents a set of stylized facts about online exporters and motivates the experiment. Section 4 describes the experiment design and main findings. Sections 5 and 6 build and estimate an empirical model of the online market. Section 7 performs counterfactual analyses. Section 8 concludes.

[4]Despite the growing importance of e-commerce in international trade, empirical work on the online setting has remained scarce and has so far primarily focused on patterns of online trade and the role of geographic distance (Hortaçsu, Martínez-Jerez, and Douglas, 2009; Lendle et al., 2016). In this paper, we examine exporter growth dynamics in online trade and study the roles of search and information frictions. We establish a set of new stylized facts about e-commerce exporters. These facts point to new trade models that extend the standard heterogeneous firm and trade framework to incorporate important features of the online marketplace. Our work, by taking into account information frictions in the online market, relates to the extensive literature on online reputation mechanisms. We refer interested readers to Tadelis (2016) for an excellent review.

2 Empirical Setting and Data

In this section, we introduce the setting of the study, the market of children's t-shirts on AliExpress, and describe the data collection.

2.1 The Market of Children's T-shirts on AliExpress

AliExpress, a subsidiary of Alibaba, was founded in April 2010 to specialize in international trade. As a global leading platform for cross-border B2C trade, AliExpress serves over 150 million consumers from 190 countries and regions, attracting over 200 million monthly visits.[5] Over 100 million products, ranging from clothes and shoes to electronics and home appliances, and 1.1 million active sellers, primarily retailers located in China, are listed on the platform.[6] Most sellers on the platform are retailers rather than manufacturers, and they source products from factories all over the country to export through the platform. Therefore, quality, in this context, captures firms' sourcing ability (i.e., ability to source high-quality products from manufacturers) as well as the quality of marketing and shipping services.[7]

For this study, we focus on the industry of children's t-shirts. As the largest textile and garment exporting country in the world, China accounted for over a third of the world's total textile and garment exports in 2019 (WTO, 2020). In the world of e-commerce, textile and apparel amount to 20 percent of China's total online retail, including sales on Alibaba's platforms.[8] In the category of children's t-shirts, AliExpress hosted over 1,800 sellers with close to 20,000 listings by 2017, fostering a vibrant market environment with substantial entry and growth dynamics. Therefore, it provides an ideal setting to study exporter dynamics. In addition, the t-shirt product category features well-specified quality dimensions, making it possible to construct *direct* quality measures to study quality-size distributions and allocative efficiency.

Two features of the platform are worth highlighting. First, AliExpress does not require a sign-up fee to set up a store and list a product, thereby essentially eliminating the entry and fixed operation costs of exporting and allowing sellers, large and small, to tap into the export markets.[9] While this does help to bring many SMEs onto the platform, the lack of market selection can create important congestion in consumer search, resulting in an excessive number of firms and product offerings in the online marketplace competing for consumers' attention. The resulting welfare implication of having an

[5]Sources: `https://sell.aliexpress.com/` and `https://sell.aliexpress.com/__pc/4DYTFsSkV0.htm`.

[6]During our sample period, Aliexpress hosted only sellers from mainland China; starting in 2018, the platform also became available to sellers in Russia, Spain, Italy, Turkey, and France.

[7]While most of the sellers on the e-commerce platform are retailers instead of manufacturers, quality may still vary significantly depending on where the sellers choose to source from, high-quality versus low-quality factories, and how much quality inspection effort the sellers put in. We document this formally using detailed quality measures we collect from the study (see Section 2.2).

[8]"E-Commerce of Textile and Apparel," China Commercial Circulation Association of Textile and Apparel, 2019

[9]AliExpress charges sellers 5-8 percent of the sales revenue as a commission fee for each successful transaction. Source: https://sell.aliexpress.com/

increasing number of market participants on firms and consumers is far less clear in the presence of search and information frictions. This forms the key trade-off we seek to examine in this study.

Second, AliExpress allows us to group product listings into different *varieties*. A single *variety group* (hereafter referred to as *"group"*) may contain multiple listings, sold by different sellers, that share identical product design. This is illustrated in Figure 1. This unique feature allows us to compare listings with the same observable product attributes, thereby controlling for consumers' horizontal taste differences. We leverage this feature in our empirical analyses as described below.

2.2 Data

We collect comprehensive data from the platform, including detailed firm-product level characteristics and transaction-level sales records. We complement the platform data with objective quality measures obtained from actual purchases, direct interactions with the sellers, and third-party assessment. Below we describe the sample and the key variables used in the analyses.

(1) Census in 2017. We scraped all product listings in the industry of children's t-shirt in June 2017. We collected all the information that a buyer can view on the listings' pages, including total cumulative orders (quantity sold), current prices, discounts (if any), ratings, buyer protection schemes (if any), and detailed product specifications. We further collected information about the stores that carry these products, including the year of opening and all other products the stores carry.

(2) Transaction Records. For each product listing, we take advantage of a unique feature of AliExpress during our sample period that allows us to keep track of a listing's most recent 6-month transaction history. For each transaction, we observe information on sales quantities, ratings, and previous buyers' countries of origin. In contrast, most existing e-commerce platforms report only customer reviews and the total volume of transactions without the full transaction history (e.g., Amazon and eBay). The availability of the real-time transaction records enables us to closely track each product listing's sales activities over time.[10]

(3) Measures of Quality Finally, we complement the platform data with a rich set of objective quality measures we collected for the study, covering quality of products measured in 8 dimensions, quality of shipping, and quality of seller service. These quality measures were collected through three channels: (i) actual purchase of the products, (ii) direct communications with the sellers, and (iii) third-party assessment. Appendix B.1 provides a detailed discussion of the quality measurement process. Table 1

[10]The transaction level data was collected twice, once in June 2017 for the universe of listings in the census data and once in August 2018 for the experimental sample (see Section 4). Each round covers six months prior to the date of the data collection. In Section 3, we use the 2017 data, together with the census data described above, to establish stylized facts about the online market. We use the August 2018 transaction history data to closely track the experimental sample and study the treatment effects, which we describe in detail in Section 4. Since the transaction data omits information on price, we further conducted a weekly data scraping from May to August 2018 for listings in the experimental sample to track price dynamics.

presents summary statistics of the various quality measures.[11]

To measure product quality, we placed actual orders of children's t-shirts on AliExpress.[12] After receiving and cataloging the orders, we worked with a large local consignment store of children's clothing in North Carolina to inspect and grade the quality of each t-shirt. The grading was done on a rich set of metrics, following standard grading criteria used in the textile and garment industry. Specifically, quality was assessed along 8 dimensions: durability, fabric softness, wrinkle test, seams (straightness and neatness), outside stray threads, inside loose stitches, pattern smoothness, and trendiness. Figure 2 Panel A shows a picture of the grading process and the criteria used. Quality along each dimension was scored on a 1 to 5 scale, with higher numbers denoting higher quality. Most of the quality metrics, except trendiness, capture vertical quality differentiation. For example, at equal prices, consumers would prefer t-shirts with more durable fabric, straighter seams and fewer loose stray threads. Exploiting the grouping function, we can further compare quality across t-shirts of the exact same design but sold by different sellers. As shown in Panel B of Figure 2, there exists considerable quality difference both across and within groups.[13]

To measure shipping quality, we recorded the date of each purchase, the date of shipment, the date of delivery, carrier name, and the condition of the package upon arrival. The information is used to construct four measures of shipping quality: (i) the time lag between order placement and shipping, (ii) the time lag between shipping and delivery, (iii) whether the package is delivered, and (iv) whether the package is damaged.

To measure service quality, we visited the homepage of each store and sent a message to the seller via the platform to inquire about a particular product. Appendix B.1 describes the messages. We rate service quality based on whether the message received a reply, the time it took to receive a reply, and whether the questions were acknowledged and properly addressed.

In Table A.2, we find all three quality indices — product, shipping and service — to be positively correlated with the online star ratings, although the correlations are relatively weak and only statistically significant for shipping and service qualities.

2.3 Summary Statistics

Table 2 summarizes the product level (Panel A) and store level (Panel B) characteristics. There are close to 20,000 product listings in the sector of children's t-shirt. The average price is $10. About

[11]To construct the quality indices, we first standardize the quality metric in each dimension and then average across all dimensions. Table A.1 decomposes the variation of the overall quality index to that explained by each individual quality metric.

[12]We placed an order on each listing in our experimental sample as well as their medium-size and superstar peers in the same variety group (see Section 4 for details on the sampling procedure).

[13]To cross-validate the quality measures, we asked the owner of the consignment store to report a bid price (willingness to pay) and a resell price for each t-shirt. Reassuringly, the objective quality metrics are strongly correlated with the subjective price evaluations.

70 percent of the listings offer free shipping, and the average shipping price to the US is \$0.50. At the store level, there are over 1,800 stores operating in this sector. Most exporters are young with an average age of two years. The average cumulative sales is 7,220 with a standard deviation of 16,618, indicating large performance heterogeneity. We observe similar patterns of performance heterogeneity at the listing level. At a given point in time, more than 40% of the listings have zero sales and the median listing has 1 order, whereas the largest listings have accumulated more than 2000 orders.

Interestingly, when we compare the market share distributions of the online export market with the traditional offline trade, we find similar concentration in sales at the top. For the offline market, we use the Chinese Customs data in 2013, the most recent year we have, and focus on firms exporting the same product category of t-shirts (HS code 6109). Table A.3 Panel A shows the market shares of the top listings in the online export market. We see that sales are heavily concentrated at the top: for example, the top 1% of the listings account for 53% of the total sales (in terms of order number) and the top 10% accounts for more than 90% of the sales. Panel B shows that at the firm level, the online distribution looks almost as skewed as the offline distribution. The top 10% of the exporters, for example, account for roughly 80% of the total sales in both the online and offline markets. This is true even when we restrict the online sample to stores that are active in the past six months: the top 10% of these active stores account for 74% of the total sales.

When it comes to life-cycle growth trajectories, the online market appears more stagnant than the offline market. For the online market, we define age as the number of years since a seller registered on AliExpress; for the offline market, we define age as the number of years since a firm started exporting in the Customs data. Figure A.1 shows that a 3-year old store on AliExpress is only, on average, 7.8 percent larger than a newborn store. In contrast, a 3-year old offline exporter is, on average, 1.5 times larger than a new offline exporter.One potential explanation for the differential growth trajectories is that firm turnover is slower online compared to offline due to the relatively low costs of operating an e-commerce business. Such inefficient exit can exacerbate the search problem: when buyers face thousands of product offerings, it is not clear who will get to grow. In the next section, we document a set of novel stylized facts about online exporters that will enable us to better understand the factors shaping firm performance and growth in e-commerce.

3 New Stylized Facts of Online Exporters

Fact 1. *Sales performance varies within identical variety groups.*

First, we exploit the unique feature of AliExpress during our study period that allowed us group product listings into different "identical-looking" varieties. Leveraging this unique feature, we first look at how sales performance varies within a single variety group. We focus on popular variety groups with more than 10 listings. As shown in Figure 3, we see that sales are quite concentrated at the top within each group. The group's superstar, defined as the listing with the highest cumulative orders within the group, accounts for about 50% of the total sales of the group; the top 10% of listings in each group captures more than 75% of the group's total orders, and the top 25% captures nearly all.

Nonetheless, looking at the distribution of the superstar sales across groups, it is also clear that this is not a case of winner-takes-all; some amount of dispersion still remains. Given that we are comparing products with the almost identical design, we are controlling for unobserved consumer horizontal taste. In a friction-less world, one may expect that the listing with the highest quality, relative to price, would win the market. The fact that some dispersion remains indicates that frictions exist in this marketplace. This raises the question of who gets to grow in the presence of these frictions. To delve more into that, we next ask who gets to become superstars.

Fact 2. *Superstars do not necessarily have the highest quality and quality only weakly predicts sales.*

We compare the quality of the superstar listings and small listings in each variety group. Superstar is defined to be the listing with the highest sales in the group and small listings are those with fewer than 5 cumulative orders. Figure 4 plots the distribution of quality difference between the group superstar and the average of the small listings in each group. We observe a substantial fraction below zero: superstars actually have lower quality than the small listings in 45% of the variety groups we sampled. Consistent with this, Figure 5 looks at how quality predicts sales. We see that the average market share of a listing only weakly increases with quality. The difference is not significant except at the top.

These observations are indicative of potential misallocation in this market due to search and information frictions. Intuitively, search friction introduces a random component in firm growth due to the consumer sampling process. When firms' intrinsic quality is not perfectly observed, such friction can further slow down the resolution of the information problem and hinder market allocation towards better firms. The result highlights the important interaction between search and information frictions and underscores the potential sources of market misallocation.

It is worth noting that the evidence is only suggestive because we have to take into account price differences. Interestingly, we find that superstars do not always charge the lowest price: within an identical variety group, the listing with the highest sales only charges the lowest price for 14% of the time. On the other hand, we do observe a positive relationship between price and quality, which corroborates our quality measures but could mean that this relatively flat relationship between quality and sales can be partly driven by price. Therefore, to isolate the role of information friction and quantify the degree of misallocation, we rely on a structural model in Section 5.

Fact 3. *On average, it takes 64 days for the first order to arrive; after that, subsequent orders arrive much faster.*

Finally, we delve more into the growth dynamics and examine how superstars emerge. Using the transaction-level data over a period of six months from January to June 2017, we explore the dynamics of order arrivals. Figure 6 plots the number of days it takes to receive the n-th order. Panel A shows the order arrival dynamics for the full unbalanced sample of all listings; Panel B restricts to listings that accumulated more than 10 orders during the six-month period. A striking pattern emerges: on average it takes 44-64 days for the first order to arrive; however, conditioning on having one order, subsequent orders arrive much faster. For example, on average the second order arrives 3-5 days after the first order, and the third order arrives 3-4 days after the second. Table A.4 regresses the dummy of receiving an order in a given week on log of past cumulative orders of a product listing, with and without store fixed effect. The results show that past sales influence future sales: firms with larger past sales, hence higher visibility, have an advantage in overcoming the search friction and generating future orders. However, the same force can lead to allocative inefficiency in the presence of information friction. In particular, the demand accumulation process implies that initial demand shocks (or "luck") can have a persistent impact on firms' long-run growth, as opposed to firm fundamentals, such as quality or productivity. Over time, firm performance diverges; market allocation and consumer welfare depend crucially on the interactions of these demand-side forces.

However, a key empirical challenge of identifying the role of the demand-side effects is to control for unobserved supply-side actions. For example, it could be that after some initial period of preparation, sellers start to invest in some costly actions, such as paying for advertising or participating in sales and promotion events organized by the platform, which then lead to the first order as well as subsequent orders. From the observational data, it is difficult to tease apart the demand- and supply-side channels. This motivates us to conduct an experiment to identify the role of demand.

4 Experiment and Findings

To demonstrate the role of demand in helping firms to overcome search and information frictions in e-commerce, we conduct an experiment in which we generate exogenous demand and information shocks to a set of small sellers via randomly placed online orders and reviews. We describe the experiment design and present the main findings below.

4.1 Experiment Design

Sampling: Prior to the start of the experiment, a baseline data collection was conducted in May 2018 that covers the majority of the product listings in the sector of children's t-shirts on AliExpress.[14]

[14]Unfortunately, due to an important redesign of the platform algorithm, only the first 100 pages were accessible. This excludes a large number of small and new listings. Therefore, we rely on the 2017 data for establishing the stylized facts in Section 2. We use the May 2018 data for selecting the experimental sample and constructing baseline controls. Table A.5

We group the product listings into 4,640 distinct variety groups using the grouping function described above. We focus on the "popular" varieties with greater than 100 orders, aggregated across all stores, and sold by at least two "small" stores with fewer than 5 cumulative orders. This screening procedure enables us to create a treatment and control group of "identical" product listings. In total, 133 varieties satisfy the above criteria, containing 1,265 product listings from 638 stores.

Randomization: Of the 1,265 product listings, 790 are small listings with fewer than 5 orders. We randomize the 790 small listings into three groups of different order and review treatments: a control group C without any order and review treatment, T1 which receives 1 order randomly generated by the research team and a star rating, and T2, which, in addition to receiving an order and a star rating, further receives a detailed review on product and shipping quality.

Given that ratings are highly inflated on AliExpress (out of the 737,000 reviews we observe over a 6-month window, 87% are five stars), for all the treatment groups we leave a five-star rating to the order unless there is any obvious quality defect or shipping problem. This is to mimic the behavior of actual buyers. To generate the contents of the shipping and product reviews, we use the Latent Dirichlet Allocation topic model in natural language processing to analyze past reviews and construct the review messages based on the identified key words. Appendix B.2 describes the reviews in detail.

The difference between T1 and C identifies the impact of demand. The difference between T1 and T2 identifies any additional impact of alleviating information frictions. To allow for comparisons across otherwise "identical" listings, we leverage the grouping function and stratify the randomization by variety group. For varieties sold by two small sellers (and other big sellers), we assign 1/2 to control and 1/2 to treatment. The latter is randomly split into T1 and T2 with equal probabilities. For varieties sold by more than two small sellers (and other big sellers), we assign 1/3 to each of C, T1, and T2. This randomization procedure is powered to identify the impact of receiving an order, followed by the impact of reviews. In the end, we have 303 listings in C, 259 in T1, and 228 in T2. Table A.6 presents the balance checks and shows that the randomization was balanced across baseline characteristics.[15]

4.2 Treatment Effects of Demand and Information Shocks

We define a dummy variable "Order", which equals to 1 if a listing received the order treatment regardless of the review treatment (i.e., in T1 or T2). Figure A.2 plots the distribution of cumulative

presents the same summary statistics at the listing, store, and group level for the baseline sample. Compared to Table 2, the number of observations is much smaller, due to the page limit issue described above. In particular, the sample captures some but not all of the small listings with close-to-zero sales.

[15]We leverage the experiment design to collect information on quality. Measuring product and shipping quality involves making actual purchases. In addition to the 487 small listings in the two treatment groups, we also purchased from the largest listings in the 133 variety groups as well as all the medium-size listings with cumulative orders between 6 and 50 in the same variety groups. This allows us to examine the relationship between quality and size. For service quality, we reached out to all 638 stores in the 133 variety groups and directly communicated with the sellers via the platform. For those with multiple listings included in the 133 groups, we randomly selected one listing for inquiry.

net orders (subtracting our own order) 3 months after the intervention. We see a small shift to the right among the treated listings, especially from the 0 to 1 margin. Overall, most listings remain small except for a few outliers that managed to grow.[16]

Next, we estimate the following regression to examine the impact of the order and review treatments on weekly cumulative orders after the initial order treatment:

$$\text{WeeklyOrders}_{it} = \alpha + \beta\text{Order}_i + \gamma_1\text{Review}_i \times \text{PostReview}_t + \lambda_t + \nu_{g(i)} + \epsilon_{it} \tag{1}$$

where the dependent variable is the total number of orders (excluding our own order) for listing i in week t. Order is a dummy variable for receiving the order treatment (which equals 1 for T1 and T2). Review is an indicator for receiving additional shipping and product reviews (T2). PostReview is a time dummy variable that equals 1 after the reviews were provided in week 7. The specification leverages the panel structure of our data since the reviews were only given upon receiving the orders.[17] λ_t and $\nu_{g(i)}$ are week and group fixed effects. In addition, all regressions control for baseline sales, both at the store level and the listing level. Standard errors are clustered at the listing level.

The results are shown in Table 3. We include baseline controls of cumulative orders at the store level and product level in the regressions. Results without these baseline controls are very similar and are shown in Table A.7. We see that the order treatment has a small but significantly positive impact on subsequent orders. This demonstrates that indeed a key channel for firms to grow in the online marketplace is by accumulating demand. On the other hand, the impact of the reviews are insignificant, suggesting that the online reputation mechanism may not function effectively in the presence of large search friction. Intuitively, reviews only matter when consumers click and visit a seller's product listing page, which is a rare event for small sellers due to their low visibility.

Table 4 examines the dynamic effects of the order treatment and shows that the effect is salient in the short run (i.e., the first month) but decays quickly afterwards. This is consistent with the fact that only a small number of sellers are able to take advantage of the short-term boost in visibility generated from the order treatment to overcome the initial hurdle of growth.[18] Quantile regression results in Table 5 further show that the impact of the order treatment concentrates in the very top quantiles while the majority of the listings experience no significant impact, consistent with the endline cumulative net orders distribution in Figure A.2. At the same time, heterogeneous treatment effect analyses in Table A.9 show that these rising new stars are not necessarily those with high quality.[19] This echoes the

[16]We focus on the impact on orders instead of revenue since we observe very little price adjustment during the study period. In the 13 weeks following the initial treatment, only 6.5% of the listings have experienced any price adjustment.

[17]Most of the orders, 801 out 826, arrived within the first 7 weeks. 2 orders arrived later and 23 orders went missing. We left the online reviews in week 7 after the initial order placement when we had received majority of the orders.

[18]In Table A.8, we examine the treatment effect on listings' relative ranking and find that receiving one order indeed leads to a small, short-term improvement in listing visibility.

[19]Here we interact the treatment variable with service quality and listing ratings because product quality and shipping

stylized fact in Section 2 that quality does not strongly predict growth in this market. In this market environment, search and information frictions combined can make it difficult for high-quality sellers to stand out.[20]

All together, the experimental findings are consistent with the presence of search and information frictions, and show that in such an environment accumulating initial demand acts as a crucial force in shaping firms' subsequent growth. Having said that, the size of the estimated average treatment effect ranges from 0.11 to 0.25, as shown in Table 5. The magnitude is much smaller than 1, which explains why individual sellers would not replicate the order treatment themselves and suggests that the demand-side frictions cannot be easily overcome by individual sellers' private efforts.[21] Next, motivated by the reduced form evidence, we build a structural model of the online market incorporating these realistic frictions of the market.

5 Model

Our model focuses on demand-side mechanisms to highlight the role of the search friction due to limited sample search and the information friction due to noisy signals. Importantly, we allow the visibility of the seller to increase in its cumulative orders, reflecting the fact that products sold by larger sellers are often positioned more saliently on the platform. We incorporate seller-side heterogeneity in both quality and cost and model sellers' pricing decisions. In this model, buyers do not directly observe quality at the point of transaction, but observe imperfect signals based on past reviews and form their beliefs. In a given period, buyers conduct a non-sequential search and randomly sample a set of sellers of different size and review history. We structurally estimate the model and perform counterfactual analysis examining the impact of initial demand, search and information frictions on firm growth, consumer welfare, and market allocation.[22]

5.1 Demand

Search

We assume that consumers conduct a fixed sample search for children's t-shirts upon their arrival.

quality are not measured for the control-group listings.

[20] We also examine the treatment effect on seller effort and business strategy. We find in Table A.10 that receiving a small order does not lead to any noticeable adjustment in pricing, shipping service, listing description (reflecting advertising effort), and introduction of new listings.

[21] In addition, the cost of manipulating orders on Aliexpress (an exclusively cross-border platform) is fairly significant and greater than that on domestic platforms. It requires recruiting people overseas and gaining access to a foreign address, foreign bank account, and foreign IP. If a buyer account or credit card is found to be repeatedly placing orders on listings carried by the same store, the account is at risk of being blocked.

[22] We have also constructed a simple baseline model that features limited sample search and flexible functional form for the visibility. Theoretical results on the limiting market share distribution can be found in the Appendix C.

Their search sample size K is exogenously given and will be estimated later. The search procedure is a weighted random sampling without replacement, and the weight for seller $i \in I$ depends on his/her relative visibility $\frac{v_i}{\sum_{j=1}^{N} v_j} \in (0,1)$. Specifically, consider a sample $\phi = \{\phi^1, \phi^2, \ldots, \phi^K\}$ of size K, where each element $\phi^k \in I, k = 1, 2, ..., K$. The probability that this sample is selected by the consumer is

$$P(\phi|\mathbf{v}) = \sum_{\mathcal{P}(\phi)} \frac{v_{\phi^{(1)}}}{\sum_{j\in I} v_j} \cdot \frac{v_{\phi^{(2)}}}{\sum_{j\in I\setminus\{\phi^{(1)}\}} v_j} \cdot \quad \cdots \quad \cdot \frac{v_{\phi^{(K)}}}{\sum_{j\in I\setminus\{\phi^{(1)},\phi^{(2)},\ldots,\phi^{(K)}\}} v_j}, \tag{2}$$

where $\mathcal{P}(\phi)$ is the set of all permutations of ϕ and $\{\phi^{(1)}, \phi^{(2)}, \ldots, \phi^{(K)}\}$ denotes one specific permutation.

We assume that each seller's visibility is increasing in their cumulative orders s_i such that $v_i = s_0 + s_i$. We allow for a basic level of visibility s_0 for sellers who have never made a sale on the platform.[23] While a more general functional form can be assumed for the visibility function $v_i = v(s_0 + s_i)$, in appendix C, we show that the curvature of $v(.)$ has important implications for the long-run limiting distribution of cumulative orders. We choose the linear functional form such that there exists a non-degenerate and non-uniform limiting market share distribution. This is consistent with our data observations.

Belief of Quality

Consumers observe the posted price p_i, the previous cumulative orders s_i, and the ratings r_i of each seller i in its search sample. However, they do not observe the true quality q_i of seller i. If the seller has never made any sale on the platform, i.e. $s_i = 0$, then by definition, price is the only observable. We assume, in this case, consumers have a common prior belief of the seller's quality as $q_i \sim N(\mu_0, \sigma_0^2)$ where $\mu_0 = 0, \sigma_0^2 = 1$. Later we will standardize our empirical quality measures to be consistent with this assumption. For the sellers who have already made sales ($s_i > 0$), consumers will combine the sellers' previous rating information and the prior belief to form their quality expectation. We assume the Bayesian updating rule, such that

$$q_i|r_i, s_i \sim N\left(\frac{s_i r_i/\sigma^2}{1/\sigma_0^2 + s_i/\sigma^2}, \left(\frac{1}{\sigma_0^2} + \frac{s_i}{\sigma^2}\right)^{-1}\right)$$

where r_i is the *average* rating of the consumers from the previous s_i orders. As we will explain later, these ratings convey information of the true quality of seller i . σ governs the noisiness of these signals.

Purchase and Review

Consumers maximizes expected utility given their search sample ϕ, i.e.

$$\max_{i\in\phi\cup\{0\}} E[u_i|r_i, s_i] = \beta + E[q_i|r_i, s_i] - \gamma p_i + \varepsilon_i$$

[23] Appendix C shows that only the relative magnitude of s_0 and N matters for the limiting market share distribution. Therefore, we fix the number of sellers to be $N = 1000$ and only estimate s_0 as a parameter.

where ε_i is consumer's idiosyncratic preference for seller i assumed to be type I extreme value and I.I.D. We denote the no-purchase option by $i = 0$ and normalize this option to provide zero utility. If a consumer ends up with purchasing from seller i, his/her shopping experience generates a noisy signal of quality q_i. After purchase, this signal is realized and the consumer leaves a noisy review as $\tilde{r}(s_i) \sim N(q_i, \sigma^2)$. The average rating at this point becomes $r_i = \frac{1}{s_i}\sum_{n=1}^{s_i} \tilde{r}(n)$.

5.2 Supply

On the supply side, each seller is characterized with exogenous cost c_i and fundamental quality q_i. They are drawn from a random distribution upon the firm's entry into the online platform. As is often assumed in the trade and quality literature, we allow for correlation between c_i and q_i. However, we assume that neither individual sellers nor consumers are sophisticated enough to dissect the population correlation of c and q, such that there is little room to use product price as a signal for unobserved quality.[24]

Seller's Price Adjustment Problem

Since the consumer's search depends on each seller's previous cumulative orders, one might naturally think that sellers would have incentive to compete for future demand by dynamic pricing. However, in our sample, we observe very infrequent price adjustment.[25] More importantly, we do not observe systematic pattern of price increase as sellers grow their cumulative orders.

As a result, we assume that each seller has an exogenous probability of adjusting its price after a certain period of time. The frequency is directly matched to the empirical frequency of price adjustment. When sellers adjust their prices, they *do* recognize that they will be competing with a small set of rivals if they end up in a consumer's search sample. Their perceived demand is denoted as D_i. D_i depends on the rich set of public information $\mathbf{p}, \mathbf{r}, \mathbf{s}$, which are the prices, ratings, and cumulative orders of all sellers at the time of price adjustment:

$$D_i(\mathbf{p}, \mathbf{r}, \mathbf{s}) = \sum_{\phi \in \Phi_i} \tilde{P}(\phi|\mathbf{s}) \frac{\exp[Eq_i(r_i) - \gamma p_i]}{1 + \sum_{j \in \phi} \exp[Eq_j(r_j) - \gamma p_j]} \tag{3}$$

where Φ_i is all possible size K samples *that includes seller* i, ϕ is one specific realization of such a sample, and $\tilde{P}(\phi|\mathbf{s})$ is the probability that ϕ is drawn conditioning on seller i entering the search sample. Specifically,

$$\tilde{P}(\phi|\mathbf{s}) \equiv \frac{P(\phi|\mathbf{s})}{\sum_{\phi \in \Phi_i} P(\phi|\mathbf{s})}, \forall \phi \in \Phi_i.$$

In practice, this weighted summation is approximated by simulation.

[24]We found little empirical evidence of the life-cycle price dynamics for sellers, in particular, for those with higher measured quality.

[25]In our study sample, with 1265 listings, only 82 adjusted their prices during the 13 week post-treatment periods.

Given D_i, seller i solves the following problem:

$$\max_{p_i} D_i \cdot (p_i - c_i)$$

where the first order condition reads

$$p_i - c_i = -\frac{D_i(\mathbf{p}, \mathbf{r}, \mathbf{s})}{\partial D_i / \partial p_i(\mathbf{p}, \mathbf{r}, \mathbf{s})} \tag{4}$$

Given the additive structure of D_i based on the realized samples ϕ, we can easily define the key piece of demand elasticity with

$$\frac{\partial D_i}{\partial p_i}(\mathbf{p}, \mathbf{r}, \mathbf{s}) = -\gamma \sum_{\phi \in \Phi_i} \tilde{P}(\phi|\mathbf{s}) \left(\frac{\exp[Eq_i(r_i) - \gamma p_i]}{1 + \sum_{j \in \phi} \exp[Eq_j(r_j) - \gamma p_j]} \right) \left(1 - \frac{\exp[Eq_i(r_i) - \gamma p_i]}{1 + \sum_{j \in \phi} \exp[Eq_j(r_j) - \gamma p_j]} \right) \tag{5}$$

This formula makes it clear that, similar to a standard discrete choice model, a seller's own elasticity is decreasing in its market share, conditioning on being in the consumer's search sample ϕ. However, this strategic consideration becomes less pronounced as $\sum_{\phi \in \Phi_i} P(\phi|\mathbf{s})$ decreases. When the market share is highly concentrated and the probability of a small seller being chosen is close enough to zero (i.e., D_i does not depend on p_i), we assume these sellers will set their price based on a constant markup added to c_i.

Entry

Sellers enter at the same time by paying a lump sum entry cost. Upon entry, each seller gets a random draw of quality q and cost c. Sellers then set their initial prices accordingly. We can recover the entry cost from the standard free entry condition by computing the discounted future payoff of an average entrant.

6 Estimation

6.1 Parametrization and Model Identification

Our model has six structural parameters. The consumer demand depends on the constant and price coefficient in mean utility, β and γ; the review signal noise σ; the search sample size K; and the initial visibility parameter s_0. To allow for flexible correlation between each seller's quality q and cost c, we use a Gaussian Copula to model the dependence of their respective marginal distributions. The dependence is governed by parameter ρ.

Despite the richness of our data on sellers' online sales history, it provides relatively little information of the overtime variation in their cost. So we start by calibrating γ to the average price elasticity of

6.7 (in align with the estimates in Broda and Weinstein (2006)). Given γ, the rest of the structural parameters are estimated using the Method of Simulated Moments. We use the following data moments:

1. The distribution of cumulative sales for the sellers
2. The dependence of new order on cumulative orders
3. The regression coefficient of log price and the measured quality
4. The conditional distribution of cumulative orders for each measured quality segment

We simulate our model from the start until the sellers' average cumulative orders reach the level in our data (32 per listing). All the moments are jointly determined by the structural parameters in our model. However, some data moments are more informative about a specific parameter than others. The distribution of cumulative sales is tightly related to the initial visibility parameter s_0 and the search sample size K. Intuitively, a small initial visibility s_0 increases the relative importance of early orders in a seller's life cycle. The amplification effect of cumulative orders is more pronounced in this case, and it increases the skewness of market distribution. On the other hand, a larger sampling size K dampens this force by effectively allowing more sellers to compete for consumer attention in each period. The dependence of a seller's new order on cumulative orders plays a similar role in disciplining the parameters s_0 and K. Conditioning on K, the correlation between a seller's cumulative orders and measured quality identifies the review signal noise σ. If the review was very precise, then higher quality sellers would grow their orders rapidly once they end up in consumer's search sample. In contrast, a larger σ results in a flattened relationship between quality and the cumulative orders. Finally, a competing force that could result in a low correlation between cumulative order and quality is the cost-quality dependence ρ. Hence we also require our simulated data to be consistent with the observed correlation between price and quality.[26]

We bootstrap the weighting matrix using our data sample. We describe the detailed simulation and estimation procedures in Appendix D.

6.2 Estimation Results

Table 6 presents the parameter estimates with standard errors. The estimated search sample size $\hat{K}$ is 2, implying that each consumer ends his/her fixed sample search with 0.2% of all sellers in the market. Since the actual number of listings in the market is about $20,000$, each search sample consists of about 40 listings.[27] The estimate for ρ is 0.42. Given the empirical marginal distribution of cost and the

[26] This empirical strategy does rely on the fact that we take a stand on γ. We will conduct robustness checks on a broad range of γ to make sure our results are not sensitive to the specific calibrated value we used in our baseline.

[27] Our model abstracts away from multiple listings within a store and treats each listing as an independent entity. This simplification does not capture across-product spillovers within a store, which is likely to matter for large sellers but relatively less so for small sellers.

standard normal quality distribution, this translates into a coefficient of correlation between quality and cost of 0.32. The review noise σ is estimated to be 3.6. To intuitively understand the magnitude of this estimate, recall that the standard deviation of the prior belief for quality is 1. Under our estimate for σ, the standard deviation of the posterior belief is reduced only by 7.5% after one order is made. Overall, our estimate suggests that reviews are very noisy signals about sellers' quality and that the uncertainty about each seller's quality is resolved very slowly, i.e. only after a substantial amount of orders. This indicates that the reputation mechanism takes time to play a role even if a seller emerged in a consumer's choice set and successfully made a sale. As a result, the search friction, interacting with the information friction, constitutes the major hurdle of seller's initial growth. Lastly, the parameter s_0 that governs the initial visibility is estimated to be 0.41, suggesting a large visibility advantage brought by an early order. Specifically, consider at the initial stage of a market where one seller makes his/her first sales while all other sellers have made zero sales; the visibility for the former is about 3.4 (calculated as $(s_0 + 1)/s_0$) times larger than the latter.

Table 7 demonstrates how well our model matches the moments. Our model is over-identified. With essentially four parameters, we are able to match the market concentration, the dependence of new orders on cumulative orders, the correlations between price and quality, and the cumulative orders versus quality relationship all very well.

6.3 Treatment as Model Validation

In this section, we conduct simulation exercises using our estimated model to evaluate its ability of validating the experimental findings in Section 4.

Table 8 presents the model predicted treatment effects for various one-time demand shocks, with different group size of treated sellers and size of purchase orders. Recall that in our experiment, 2% of the sellers received our orders. Since the overall market is growing, we conduct the treatment in our model at the point when average cumulative orders per seller is the same as that in the data (32 t-shirts). As in the experiment, the size of the purchase is 1. It takes time for the new purchase to generate future orders for treated sellers. In our experiment, we evaluate the impact after 13 weeks of the treatment (during which period the total market orders grew by 48.7%). This number guides our choice of the number of post-treatment periods in the model to evaluate the result. In our baseline experiment simulation ($P = 2\%$, $O = 1$), we find that our model estimates result in a treatment effect of 0.058 – an average seller receiving a random purchase would grow his/her orders by around 0.06 pieces. It is slightly below but quantitatively comparable to the range of the average treatment effect shown in Table 5, which is between 0.115 and 0.257. We also show that when the size of orders increases from 1 to 2 and 5, the average treatment effect will go up more than proportionately. However, notice that they are always lower than the size of our treated purchase, which indicates that the market frictions are not easily overcome by the private effort of the sellers.

7 Counterfactual Analyses

We conduct counterfactual exercises to examine the role of information and search frictions on firm growth and consumer welfare. We evaluate potential policy interventions through the lens of our estimated structural model. The results are reported in Table 9.

7.1 The Role of Information Friction: Reducing the Review Signal Noise

In our model, since the seller's quality is unknown to the consumers, the review from past purchases is a crucial source of information spillover for subsequent consumers. Our estimate implies a quite noisy review signal, with a standard deviation of 3.6. In Panel A, Table 9, we compare our baseline with a case where we reduce σ to zero. In other words, we investigate a case that a seller-listing's true quality is immediately revealed when it accumulates its first order. We find that in this case, the cumulative order share significantly shifts towards the higher quality sellers. As illustrated in Figure 7, when we compare the blue bar (baseline) vs the purple bar ($\sigma = 0$), the share of the top quality quartile increases substantially to 55%. Since the current sales lead to more advantage in future consumer's search, the reduction in information friction allows the high quality sellers to accumulate orders much faster in this counterfactual case. As a consequence, we also find that average consumer welfare improved by 12.7%. Note that this takes into account the fact that higher quality charges higher price and highlight the sizable welfare gain from reallocation.

7.2 The Role of Initial Demand

To demonstrate the important role of initial demand in determining the market outcomes and welfare in the presence of search and information frictions, we compare our baseline case in which the visibility of sellers is proportional to the sum of s_0 and cumulative sales ($v_i = s_0 + s_i$) to a case where the initial visibility of sellers is set to be proportional to their true quality[28] in the *first ten periods* and then revert to the same baseline search protocol right afterward. In the first case (our baseline), initial demand is determined purely by random formulation of consumers' search sample while in the second it is determined by seller quality. Contrasting these two cases will shed light on the role of initial demand on long-run market outcomes.

In Figure 7, we can read again the cumulative market share of sellers in each of the four quality bins, contrasting the blue bar (baseline) and the red bar (initial visibility from quality). It is remarkable that

[28] We maintain the total amount of seller initial visibility at the same level as in our baseline, while assign it to each seller i based on $\frac{exp(q_i)}{\sum_j exp(q_j)}$

only ten-period differences in demand allocation generates a persistent long-run difference in market outcomes. Panel B of Table 9 reports a few summary statistics that connect market structure to allocation and welfare. In the case of quality-based initial demand, we find that the sales weighted quality is 13.2% higher than the baseline case, resulting in a gain in expected consumer surplus of 7.4%.[29]

7.3 The Role of Search Friction: Reducing the Number of Sellers

Finally, we investigate the impact of changing the number of sellers operating in the online marketplace (for example, by raising the entry costs or the costs of maintaining an active listing). Our model estimate implies that consumers only explore 0.2% of the sellers when they make a purchase. This implies that the large number of sellers could generate congestion in the buyer search process, thereby affecting the long-run dynamics of market outcomes. Figure 7 reports an alternative scenario of 500 sellers in the yellow bar. Despite starting from exactly the same market structure, higher quality sellers are discovered sooner and obtain more orders when the total number of sellers is smaller at 500. Panel C of Table 9 reports that the top quartile quality seller gains 11.2% cumulative market share when the number of sellers is reduced from 1000 to 500. As a result, the sales-weighted quality is 22% higher and the expected consumer surplus increases by 10.6%. This counterfactual exercise speaks to policies targeting at the creation of new marketplaces. For existing platforms, which already host a large number of sellers and listings, one can imagine an analogous exercise by screening out inactive seller-listings. We hope to investigate that further in the next step.

8 Conclusion

In this paper, we study exporter dynamics on global e-commerce platforms. Leveraging comprehensive data about the online businesses from AliExpress and combining that with unique objective measures of quality covering multiple product and service dimensions, we document sizable variation in firm-product quality in the online marketplace. However, we find that quality only weakly predicts firm performance and growth. Our paper highlights the role of search and information frictions in explaining the disintegration of the demand accumulation process and firm fundamentals and underscores the potential source of market misallocation in e-commerce.

Our findings speak to effective policies on facilitating small business growth via e-commerce. While global e-commerce platforms present a promising avenue for small and medium-sized enterprises in

[29] Consumer welfare is calculated as the expected sum of inclusive values from the searched sample, before and after the actual purchase decision:

$$CS = E_0 \sum_i IV(\phi_i \cup \{0\}) = E_0 \sum_i \log \sum_{j \in \phi_i \cup \{0\}} \exp(q_j - \gamma p_j),$$

where E_0 stands for the average across simulations.

developing countries to tap into the global market, simply bringing firms to these platforms may not be sufficient to generate sustained growth due to the large demand-side frictions. In fact, doing so can exacerbate the search and information problems, resulting in market misallocation. Policies should be designed to help firms, especially new businesses, to overcome the additional demand-side frictions. In the context of e-commerce, regulating entry, creating a premium market segment, and directing demand to promising newcomers could help to facilitate growth and improve the overall market efficiency.

References

Allen, Treb. 2014. "Information frictions in trade." Econometrica 82 (6):2041–2083.

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

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

Bai, Jie. 2016. "Melons as Lemons: Asymmetric Information, Consumer Learning and Seller Reputation." .

Banerjee, Abhijit Vinayak. 2013. "Microcredit under the Microscope: What Have We Learned in the Past Two Decades, and What Do We Need to Know?" Annu. Rev. Econ. 5 (1):487–519.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India*." The Quarterly Journal of Economics 128 (1).

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

Chernev, Alexander and Ryan Hamilton. 2009. "Assortment size and option attractiveness in consumer choice among retailers." Journal of Marketing Research 46 (3):410–420.

De Mel, Suresh, David McKenzie, and Christopher Woodruff. 2008. "Returns to Capital in Microenterprises: Evidence from a Field Experiment." The Quarterly Journal of Economics 123 (4):1329–1372.

Drinea, Eleni, Alan Frieze, Michael Mitzenmacher, and Michael Mitzenmacher. 2002. "Balls and bins models with feedback." In Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, 308–315.

Eaton, Jonathan, David Jinkins, James Tybout, and Daniel Yi Xu. 2016. "Two-sided Search in International Markets." In 2016 Annual Meeting of the Society for Economic Dynamics.

Fitzgerald, Doireann, Stefanie Haller, and Yaniv Yedid-Levi. 2020. "How Exporters Grow." .

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

Hansman, Christopher, Jonas Hjort, Gianmarco León-Ciliotta, and Matthieu Teachout. 2020. "Vertical integration, supplier behavior, and quality upgrading among exporters." Journal of Political Economy 128 (9):3570–3625.

Hortaçsu, Ali, F Martínez-Jerez, and Jason Douglas. 2009. "The geography of trade in online transactions: Evidence from eBay and mercadolibre." American Economic Journal: Microeconomics 1 (1):53–74.

Khandelwal, Amit. 2010. "The long and short (of) quality ladders." The Review of Economic Studies 77 (4):1450–1476.

Kugler, Maurice and Eric Verhoogen. 2012. "Prices, Plant Size, and Product Quality." The Review of Economic Studies 79 (1):307–339.

Lendle, Andreas, Marcelo Olarreaga, Simon Schropp, and Pierre-Louis Vézina. 2016. "There goes gravity: eBay and the death of distance." The Economic Journal 126 (591):406–441.

Lipsman, A. 2019. "Global Ecommerce 2019: Ecommerce Continues Strong Gains Amid Global Economic Uncertainty. eMarketer." Saatavissa (viitattu 8.4. 2020): https://www. emarketer. com/content/global-ecommerce-2019 .

Macchiavello, Rocco and Ameet Morjaria. 2015. "The value of relationships: evidence from a supply shock to Kenyan rose exports." American Economic Review 105 (9):2911–45.

McShane, Blakeley B and Ulf Böckenholt. 2018. "Multilevel multivariate meta-analysis with application to choice overload." psychometrika 83 (1):255–271.

Mitzenmacher, Michael, Roberto Oliveira, and Joel Spencer. 2004. "A scaling result for explosive processes." the electronic journal of combinatorics 11 (1):R31.

Pitney Bowes, Inc. 2016. "2016 Pitney Bowes Global Online Shopping Study." http://www.pitneybowes.com/us/global-ecommerce/case-studies/the-growing-importance-of-international-ecommerce.html. Accessed: 2020-07-20.

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

Ruhl, Kim J and Jonathan L Willis. 2017. "New exporter dynamics." International Economic Review 58 (3):703–726.

Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M Todd. 2010. "Can there ever be too many options? A meta-analytic review of choice overload." Journal of consumer research 37 (3):409–425.

Startz, Meredith. 2018. "The value of face-to-face: Search and contracting problems in Nigerian trade." .

Steinwender, Claudia. 2018. "Real effects of information frictions: When the states and the kingdom became united." American Economic Review 108 (3):657–96.

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

UNCTAD, Torbjörn, Fredriksson. 2016. "Unlocking the Potential of E-Commerce for Developing Countries." .

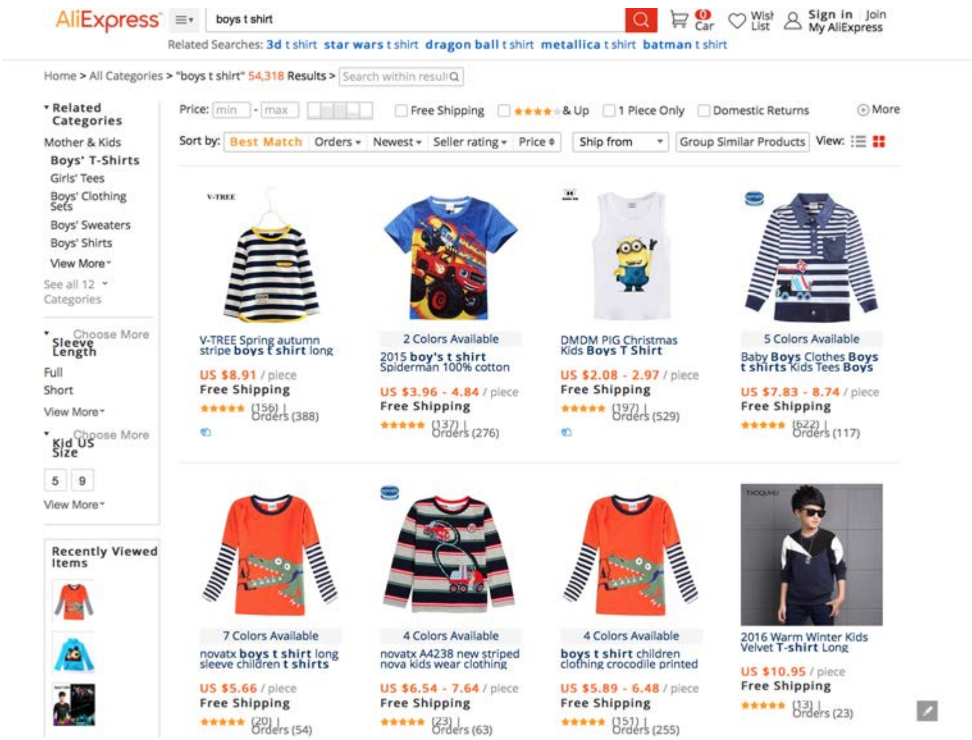
Verhoogen, Eric. 2020. "Firm-Level Upgrading in Developing Countries." Tech. rep.

Verhoogen, Eric A. 2008. "Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector." The Quarterly Journal of Economics 123 (2):489–530.

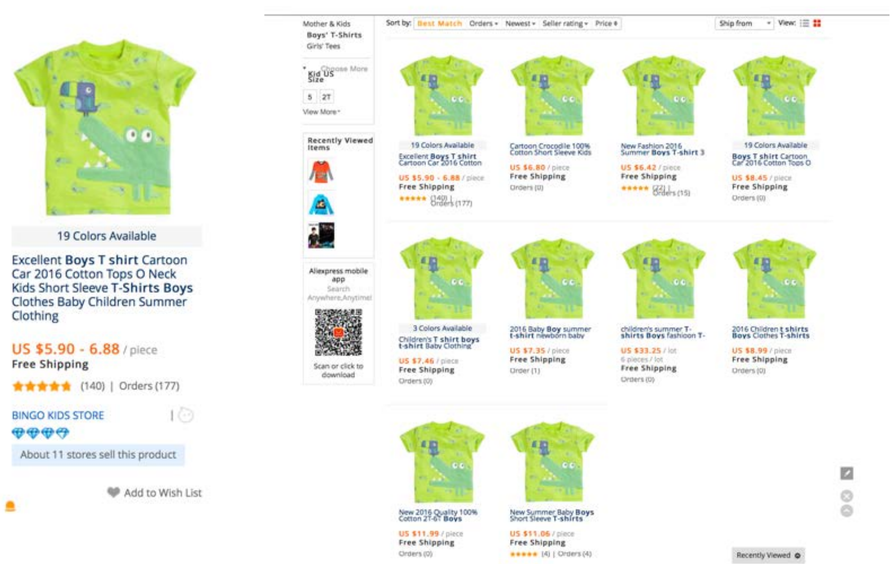
WTO, World Trade Organization. 2020. "World trade statistical review 2020." .

Figure 1: AliExpress: Search Results with and without Grouping

Panel A. Search Results without the Grouping Function



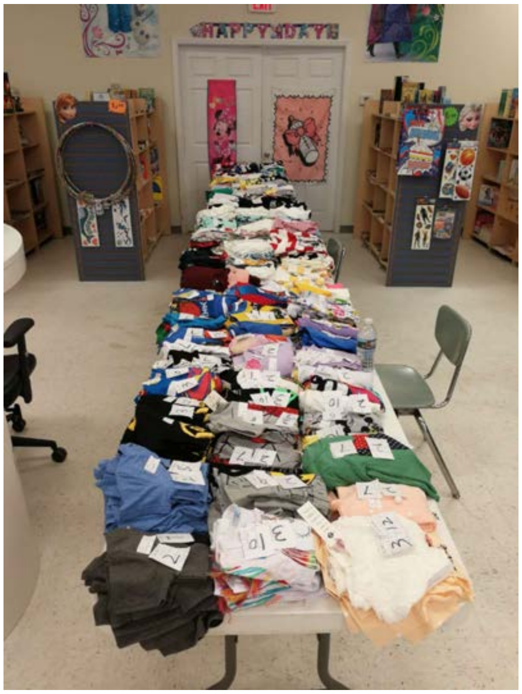
Panel B. Search Results with the Grouping Function



Note: This figure presents examples of search results on AliExpress. Panel A displays the search results using "children's t-shirts" as keywords, without applying the grouping function provided by the website. Panel B displays the same search results while applying the grouping function.

Figure 2: Quality Assessment

Panel A. Quality Assessment



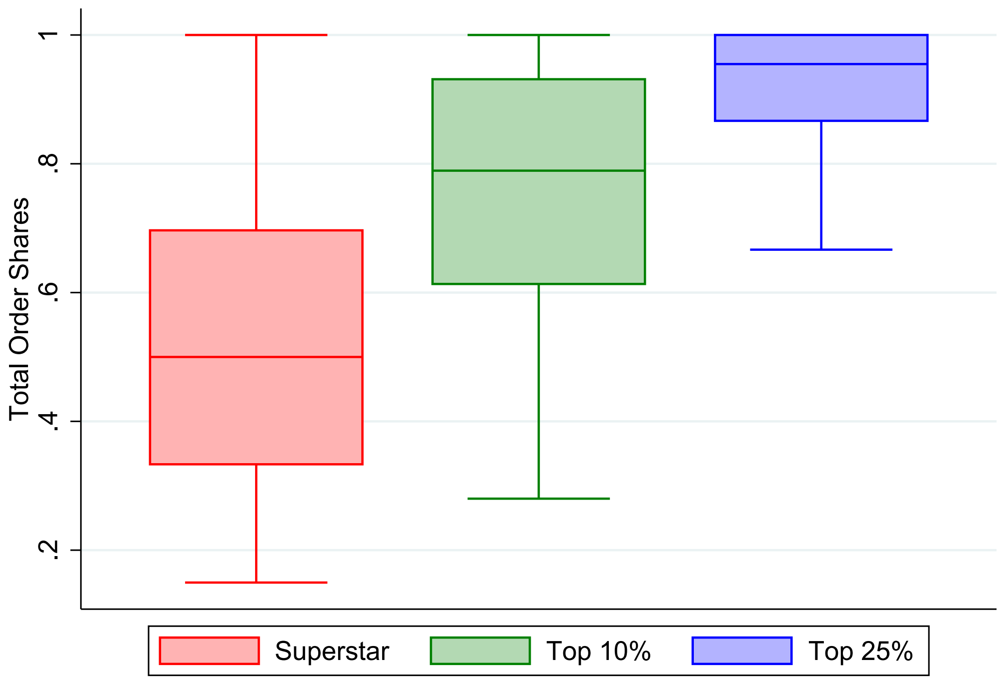
Quality Metrics:

- Obvious Quality Defect (dummy)
- Fabric/Materials (1-5 Rating) :
  - ✓ Durability/ Strength(tightly woven?)
  - ✓ Softness
  - ✓ Wrinkle test
- Seam (1-5 Rating):
  - ✓ Straight and neat (e.g. armpit)
  - ✓ Outside stray threads
  - ✓ Inside multiple unnecessary/loose stitches
- Pattern Printing (1-5 Rating):
  - ✓ Smoothness
  - ✓ Trendiness (subjective)

Panel B. Variation in Quality

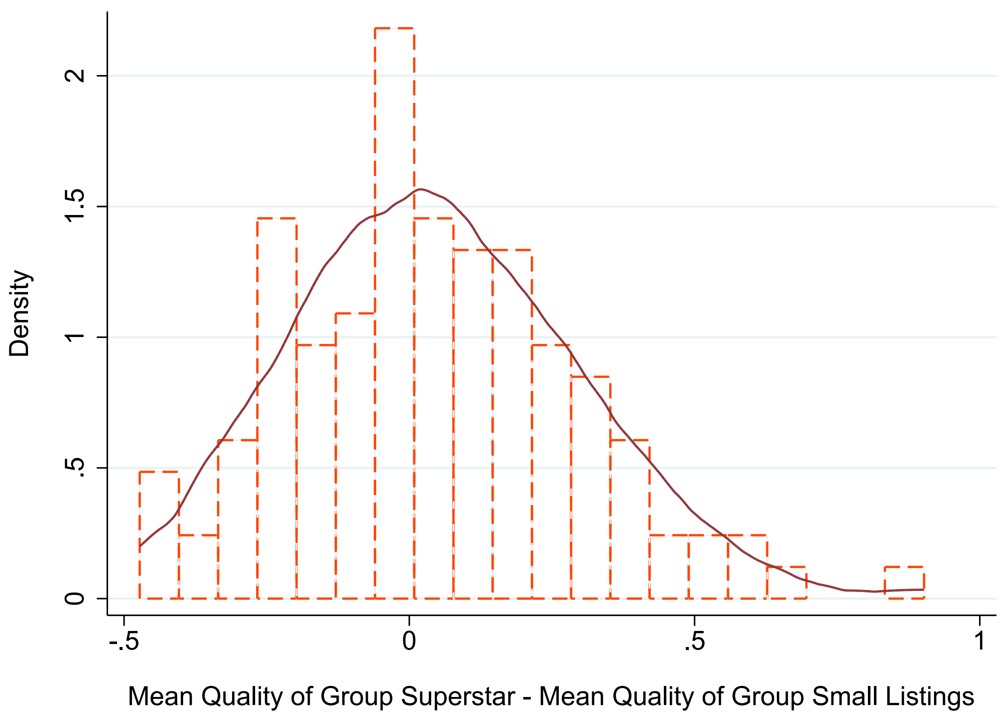
Note: Panel A displays the purchased t-shirts, sorted by groups, from our experiment; the quality assessment agent located in Durham, North Carolina; and the quality metrics used in the assessment process. Panel B shows examples of the t-shirts that receive low scores in specific quality metrics.

Figure 3: Sales Performance Within Identical Variety



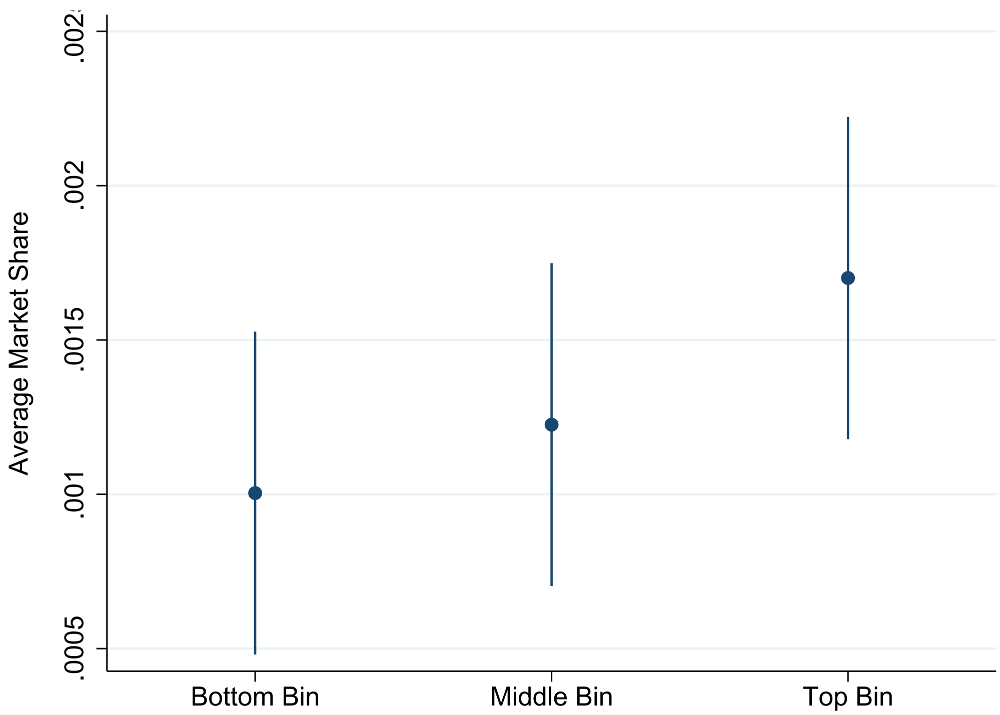
Note: This figure plots the distribution of the total share of cumulative orders for top listings across groups using the census sample in 2017 (see Section 2.2 for details on the census data). "Superstar" indicates listings that have the highest cumulative orders within its group variety. "Top 10%" ("top 25%") indicates listings that have the top 10% (25%) cumulative orders within its group variety.

Figure 4: Quality Comparison Between Group Superstar and Small Listings



Note: This figure plots the distribution of the quality differences between group superstars and group small listings using the experiment sample with quality measures (see Section 4 for details on the experiment sample). Quality is measured by the Overall Quality Index (see Section 2.2 for details on the construction of quality indices). Group superstar is defined to be the listing with the largest number of cumulative orders in each group. Small listings is defined to be the listings with fewer than 5 cumulative orders. The sample consists of the 133 groups in our experimental sample.

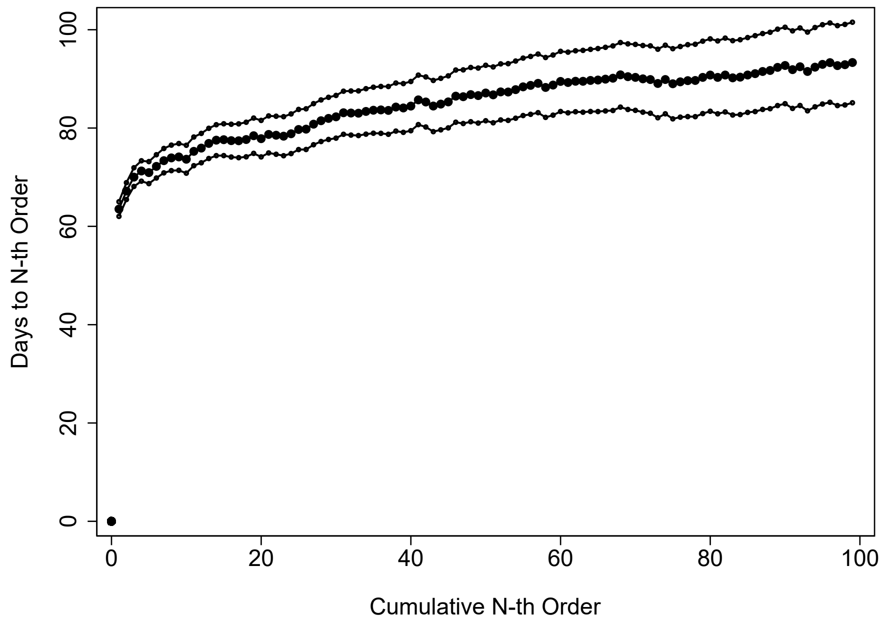
Figure 5: Average Market Share over Quality Bins



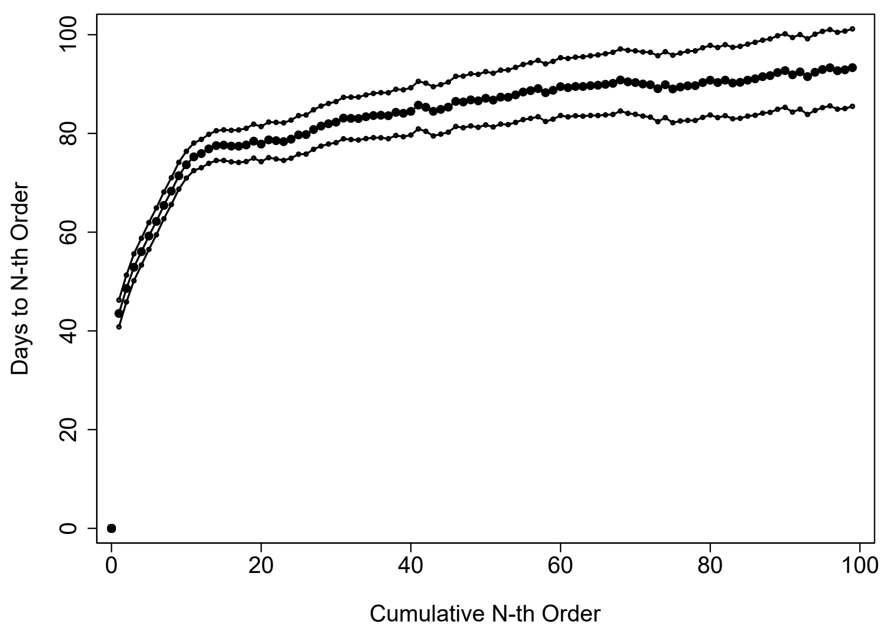
Note: This figure plots the regression coefficients and their 95% confidence intervals from regressing the listings' shares based on cumulative orders on the quality bins they belong to. The data used for the regression is the experiment sample with quality measures (see Section 4 for details on the experiment sample and Section 2.2 for the construction of quality measures).

Figure 6: Dynamics of Order Arrival

Panel A. Unbalanced Panel of All Listings

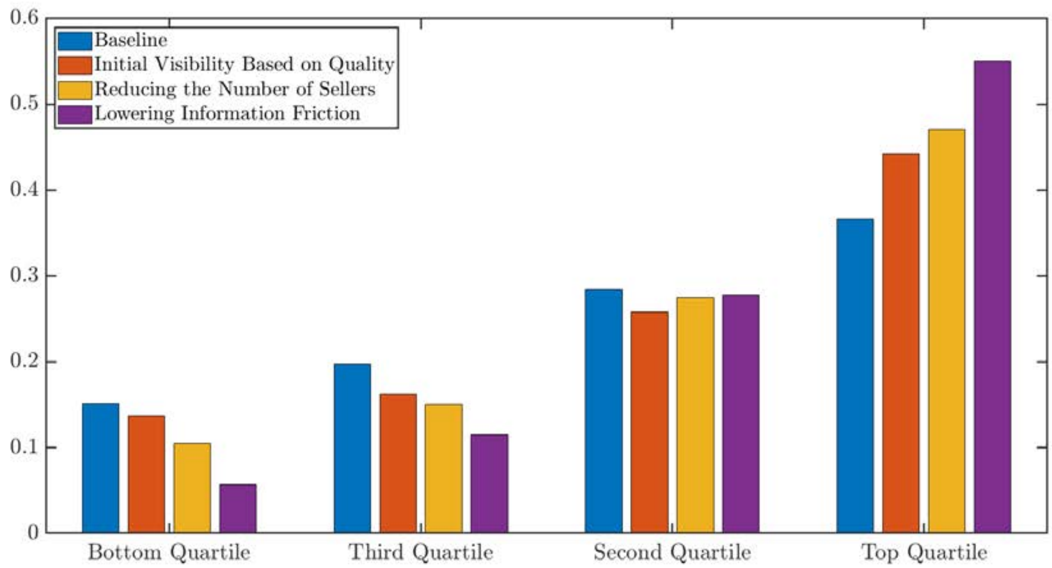


Panel B. Listings with More than 10 Cumulative Sales



Note: This figure describes the order arrival dynamics. The x-axis indicates the n-th order and the y-axis shows the number of days till receiving the n-th order. The bold line in the middle plots the average and the other two dotted lines plot the 95% confidence interval. We use the six-month transaction history data described in Section 2.2. Panel A include the full unbalanced panel of all listings appeared in the transaction data. Panel B restrict the sample to listings that accumulated more than 10 orders during the six-month period.

Figure 7: Counterfactual Market Share Distributions Over Quality Quartiles



Note: This figure plots the total shares of cumulative orders for different quality quartiles under the baseline and three counterfactual cases.

Table 1: Summary Statistics of Quality Measures

	Observations	Mean	Std Dev	Median	Intra-Group Corr	Within-Group SD	Between-Group SD
				Panel A: Product Quality			
NoObviousQualityDefect	796	.93	.26	1	.08	.25	.07
Durability	791	2.63	.79	3	.91	.24	.75
MaterialSoftness	791	3.21	.72	3	.94	.17	.71
WrinkleTest	791	3.08	.48	3	.91	.15	.46
SeamStraight	791	4.22	.47	4	.2	.42	.21
OutsideString	791	2.8	1.56	3	.4	1.22	.99
InsideString	791	.78	1.18	0	.46	.87	.8
PatternSmoothness	773	3.43	1.53	4	.87	.55	1.44
Trendiness	791	3.13	1.35	3	.91	.4	1.3
				Panel B: Service and Shipping Quality			
BuyShipTimeLag	823	3.66	3.24	3	.32	2.67	1.85
ShipDeliveryTimeLag	802	12.92	4.15	12	.19	3.73	1.83
LostPackage	820	.02	.14	0	.09	.13	.04
PackageDamage	795	0	.05	0	0	.05	.
ReplyWithinTwoDays	1258	.69	.46	1	.08	.44	.13
				Panel C: Quality Indices			
ProductQualityIndex	769	0	.41	-.02	.7	.22	.34
ShippingQualityIndex	793	.04	.43	.12	.09	.41	.13
ServiceQualityIndex	1258	0	1	.67	.08	.96	.28
OverallQualityIndex	763	.01	.29	.01	.54	.2	.22

Notes: This table reports the summary statistics of the various quality measures. Sections 2.2 and B.2 provide details on the measurement process and each of the quality metrics. Panel C reports the aggregate quality indices constructed by standardizing scores of individual quality metrics and taking their average within each quality category. The number of observations changes slightly across quality measures because some items were lost in the shipping process and in the quality assessment process.

Table 2: Summary Statistics of the Children's T-shirts Market on AliExpress

	Observations	Mean	Std Dev	Median	5th Pctile	95th Pctile
Panel A. Listing Level						
Orders	19855	32.21	199.7	1	0	91
Price (Discounted)	18681	7.63	19.65	5.85	2.62	14.22
Revenue (Discounted)	18681	164.48	1040.85	7.02	0	454.1
Total Feedback	19855	26.49	195.67	0	0	66
Star Rating	8960	3.87	1.96	4.9	0	5
Free Shipping Indicator	19855	.72	.45	1	0	1
Shipping Cost to US	19854	.51	4.3	0	0	2.61
Panel B. Store Level						
Num of Variety	1816	159.42	193.88	93	19	514
Total Orders	1816	7219.52	16617.96	1923.5	73	32847
Total Revenue (Discounted)	1816	35687.67	77427.83	14072.43	628.55	144612.8
Total Revenue (Not Discounted)	1816	46999.66	95827.14	17750.11	822.73	189711.1
T-shirts Orders	1816	343.21	1240.35	14	0	1814
T-shirts Revenue (Discounted)	1816	1645.85	5788.11	83.59	0	8622.65
Zero T-shirts Sales Indicator	1816	.2	.4	0	0	1
Age	1816	2.32	1.76	2	0	5
Total Feedback	1816	6042.97	11254.08	1651	22	29853
Perc of Positive Feedback	1803	.97	.03	.98	.94	1
Description	1771	4.68	.15	4.7	4.5	4.8
Communication	1771	4.69	.15	4.7	4.5	4.8
Shipping Speed	1771	4.57	.18	4.6	4.3	4.8
Num of Positive Feedback	1816	6241.28	11652.19	1686.5	23	30954
Num of Neutral Feedback	1816	227.32	483.12	51	0	1252
Num of Negative Feedback	1816	191.17	426.18	41	0	1067
Num of Negative Feedback	1805	.97	.02	.98	.94	1

Note: This table reports the summary statistics for the census data in 2017 of the children's t-shirts market on AliExpress. See Section 2.2 for details on the census data. Panel A reports the summary statistics for all children' t-shirts listings on Aliexpress. Panel B reports the summary statistics for the stores selling these listings.

Table 3: Treatment Effects of Order and Review

	All Destinations		English-speaking Countries		United States	
Order	0.024	0.027*	0.015***	0.016***	0.017***	0.018***
	(0.020)	(0.014)	(0.005)	(0.005)	(0.003)	(0.003)
ReviewXPostReview	0.001	-0.017	0.020	0.016	0.016	0.014
	(0.023)	(0.027)	(0.019)	(0.017)	(0.017)	(0.016)
Observations	10270	10270	10270	10270	10270	10270
Group FE	No	Yes	No	Yes	No	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Clustered SE at listing level	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the treatment effects of the experimentally generated orders and reviews. Section 4 provides more details on the sample and procedures of the experiment. The dependent variable is the weekly number of orders, calculated using the transaction data collected in August 2018. The baseline controls include the baseline total number of cumulative orders of the store and of the particular product listing. "Order" is a dummy variable that equals one for all products in the treatment groups (T1 and T2) and zero for the control group. "Review" is a dummy that equals one for all products in T2, where we place one order and leave a review on shipping and product quality. "PostReview" is a dummy that equals one for the weeks after the reviews were given (i.e., from week 7 onward). Standard errors are in the parentheses. *** indicates significance at 0.01 level, ** 0.5, * 0.1.

Table 4: Dynamic Treatment Effects

	All Destinations		English-speaking Countries		United States	
OrderXMonth1	0.063***	0.064***	0.037***	0.037***	0.038***	0.038***
	(0.024)	(0.021)	(0.008)	(0.008)	(0.007)	(0.007)
OrderXMonth2	0.019	0.019	0.007	0.008	0.008**	0.008**
	(0.032)	(0.027)	(0.010)	(0.009)	(0.003)	(0.004)
OrderXMonth3	0.009	0.009	0.015*	0.015**	0.012*	0.013**
	(0.020)	(0.017)	(0.008)	(0.008)	(0.007)	(0.006)
OrderXMonth4	-0.044	-0.044	-0.009	-0.008	0.011	0.011
	(0.031)	(0.029)	(0.018)	(0.018)	(0.007)	(0.007)
Observations	10270	10270	10270	10270	10270	10270
Group FE	No	Yes	No	Yes	No	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Clustered SE at listing level	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the dynamic treatment effects of the experimentally generated orders and reviews. The dependent variable is the weekly number of orders, calculated using the transaction data collected in August 2018. The baseline controls include the baseline total number of cumulative orders of the store and of the particular product listing. "MonthX" is a dummy variable that equals one for the X-th month after treatment. Standard errors are in the parentheses. *** implies significance at 0.01 level, ** 0.5, * 0.1.

Table 5: Average and Quantile Treatment Effects Measured at the Endline

	(1) Average	(2) 10th %	(3) 50th %	(4) 90th %	(5) 95th %	(6) 99th %
Orders from the US	0.248***	0.010	0.032*	0.673***	0.931***	3.856
	(0.063)	(0.019)	(0.019)	(0.198)	(0.263)	(2.447)
Orders from English-Speaking Countries	0.190**	0.020	0.050**	0.637***	1.080**	3.625*
	(0.093)	(0.025)	(0.022)	(0.239)	(0.464)	(1.962)
Orders from All Countries	0.110	0.096	0.149*	0.593	0.977	3.394
	(0.308)	(0.110)	(0.077)	(0.652)	(1.216)	(4.973)
Observations	790	790	790	790	790	790
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the average and quantile treatment effects of the experimentally generated orders and reviews. Each cell in the table reports a regression coefficient. The dependent variable is the endline number of cumulative orders, calculated using the transaction data collected in August 2018. The independent variable is the order treatment dummy that equals one for all products in the treatment groups (T1 and T2) and zero for the control group. The baseline controls include the baseline total number of cumulative orders of the store and of the particular product listing. Column 1 reports the average treatment effect, and Columns 2 to 6 report the quantile treatment effects. Standard errors are in the parentheses. *** indicates significance at 0.01 level, ** 0.5, * 0.1.

Table 6: Estimated Parameters of the Empirical Model

Parameters	s_0	K	σ	ρ	γ	β
Value	0.41	2	3.6	0.42	1.2	4.2
	(0.004)	(0.119)	(0.007)	(0.014)	(0.002)	(0.007)

Note: s_0 governs the initial visibility; σ is the review noise; ρ is the parameter that maps to the correlation between cost and quality; K is the search sample size; and the price coefficient γ is calibrated by choosing a reasonable average markup; β is calibrated to match outside option share (see discussion in Section 5). Standard errors are reported in the parentheses.

Table 7: Matching Moments

Moments	**Data**	**Model**
Top 1% cumulative revenue share	0.424	0.394
Top 5% cumulative revenue share	0.742	0.661
Top 10% cumulative revenue share	0.841	0.775
Top 25% cumulative revenue share	0.939	0.906
Top 50% cumulative revenue share	0.983	0.974
Dependence of new order on cumulative orders	0.023	0.066
Cumulative orders share: 1st quality bin	0.434	0.460
Cumulative orders share: 2nd quality bin	0.311	0.322
Reg. coef. of log price and quality	0.125	0.130

Note: This table reports the data moments and the model moments evaluated at the parameter estimates. See Section 5 for more discussion on the choice of moments.

Table 8: Model Validation Using the Experiment

Percent of Sellers Purchased	Size of Purchase	Average Effect on Sales: Treated - Control
P	O	$\Delta M = 48.7\%$
2	**1**	**0.058**
2	2	0.166
2	5	0.465

Note: This table shows the simulated treatment effect based on the estimated model. The first two columns are the coverage and size of the simulated treatment, and the last column reports the increase in orders averaged over treated sellers after the total number of cumulative orders in the market increases by 48.7%.

Table 9: Counterfactual Analyses

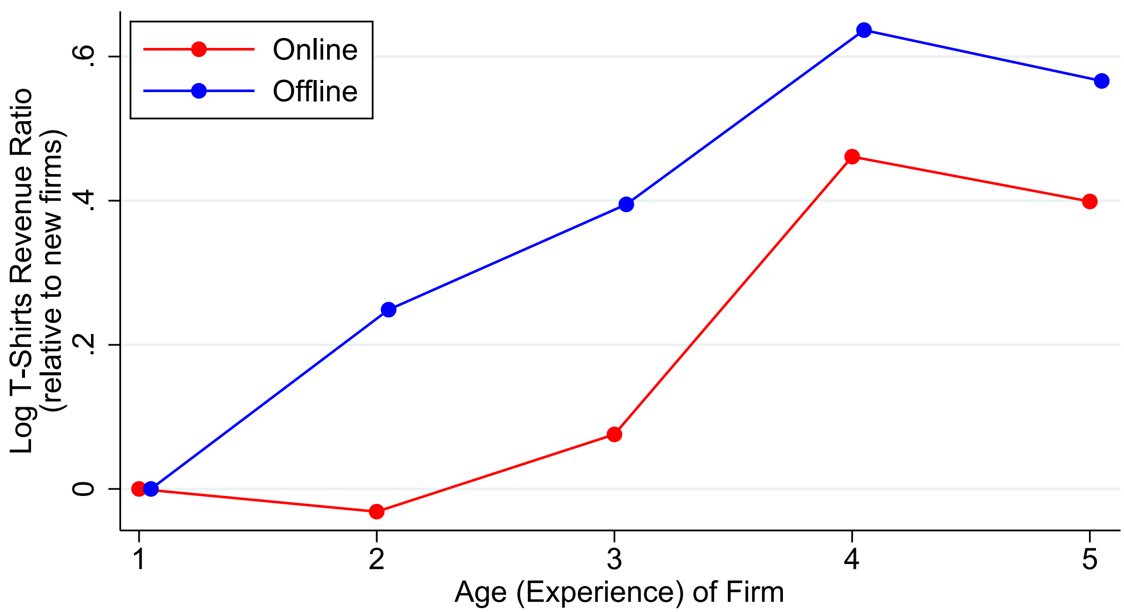
	Total Share for Top Quality Quartile	Sales-Weighted Quality	Average Consumer Surplus
		Panel A: Information Friction	
Baseline ($\sigma = 3.6$)	0.367	0.252	100.0
$\sigma = 0$	0.550	0.681	112.7
		Panel B: Search and Initial Luck	
Baseline Search Protocol	0.367	0.252	100.0
Quality-Based Search	0.443	0.384	107.4
		Panel C: Number of Sellers	
Baseline (1000 Sellers)	0.367	0.252	100.0
500 Sellers	0.471	0.474	110.6

Note: This table reports the results of several counterfactual analyses using the estimated model. Panel A compares market outcomes using sales-based versus quality-based initial visibility. Specifically, in the former case, visibility is proportional to the sum of s_0 and sales. In the latter case, visibility is proportional to exponential quality in the first period and revert to be proportional to the sum of s_0 and sales afterward. Panel B compares market outcomes with different numbers of sellers under the sales-based search protocol. Panel C compares the long-run average market outcomes in the baseline case with 1000 sellers versus in a premium market where sellers with zero sale are screened out.

Appendices. For Online Publication Only

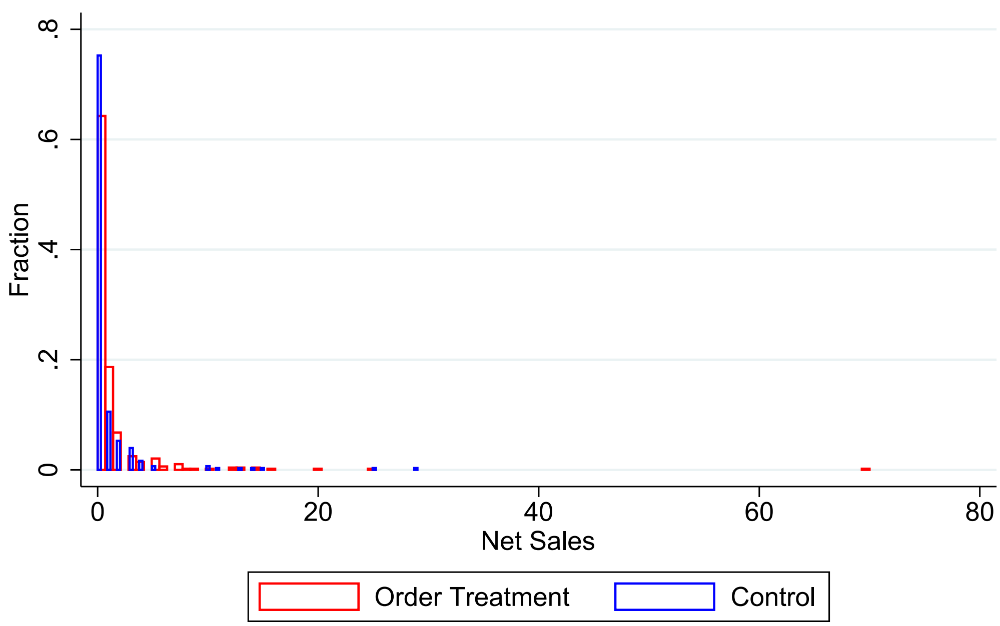
A Figures and Tables

Figure A.1: Life Cycle Growth Dynamics



Note: This figure plots the life cycle dynamics of online and offline exporters. For the online exporting firms, we use the census data collected in 2017 (see Section 2.2 for more details on the census data). For the offline exporting firms, we use the Chinese Customs data in 2013 for the corresponding HS category of t-shirts (code 6109). We calculate firm sizes as their revenue from selling t-shirts.

Figure A.2: Endline Net Sales Distribution



Note: This figure plots the endline distribution of cumulative orders (net of our own orders) for the listings in the experiment sample. The blue bars indicate the control group; the red bars indicate the treatment groups, i.e. T1 and T2.

Table A.1: Decomposition of the Overall Quality Index

Quality Metrics	Explained R^2
OverallQualityIndex	**100**
ProductQualityIndex	**76.0**
NoObviousQualityDefect	9.3
Durability	13.5
MaterialSoftness	8.8
WrinkleTest	7.1
SeamsSraight	6.6
OutsideString	8.3
InsideString	8.4
PatternSmoothness	9.7
Trendiness	4.3
ShippingQualityIndex	**18.2**
BuyShipTimeLag	3.4
LostPackage	0.0
NoPackageDamage	8.0
ShipDeliveryTimeLag	6.8
ServiceQualityIndex	**5.8**
ReplyWithinTwoDays	5.8

Note: This table decomposes the variation of the overall quality index to that explained by each individual quality subindices and metrics. For the subindices (i.e. ProductQualityIndex, ServiceQualityIndex, and ShippingQualityIndex), the Shapley value is reported. For other metrics, the Owen value is reported. We use the quality measures collected for the experiment sample. See Section 2.2 for a detailed description of the quality collection process.

Table A.2: Correlation between Quality and Star Rating

	Dependent: Star Rating					
	(1)	(2)	(3)	(4)	(5)	(6)
ProductQualityIndex	0.029	0.170				
	(0.048)	(0.114)				
ShippingQualityIndex			0.082*	0.098*		
			(0.044)	(0.055)		
ServiceQualityIndex					0.034**	0.036*
					(0.017)	(0.020)
Constant	4.803***	4.804***	4.795***	4.793***	4.793***	4.793***
	(0.019)	(0.020)	(0.019)	(0.019)	(0.016)	(0.017)
Observations	409	409	422	422	624	624
Rsquare	0.001	0.317	0.008	0.319	0.006	0.210
Group FE	No	Yes	No	Yes	No	Yes

Note: This table presents results from regressing listing star ratings on their quality indices. We use the experiment sample with detailed quality measures. See Section 4 for more details on the experiment sample and Table 1 for the summary statistics of quality measures. Standard errors are in the parentheses. *** indicates significance at 0.01 level, ** 0.5, * 0.1.

Table A.3: Market Share Distributions at the Top

	Number of Exporters at the Top		Market Share of Exporters at the Top	
	Online	Offline	Online	Offline
		Panel A. Listing Level		
Top 1%	198	.	52.5	.
Top 5%	992	.	81.3	.
Top 10%	1985	.	89.8	.
Top 25%	4963	.	97.3	.
Top 50%	9927	.	99.9	.
		Panel B. Firm Level		
Top 1%	18	120	27.7	38.2
Top 5%	90	602	62.7	65.2
Top 10%	181	1205	79.7	78.4
Top 25%	453	3012	95.0	92.7
Top 50%	907	6025	99.5	98.6
Total Num of Listings	19855	.	.	
Total Num of Firms	1816	12052	.	.
Total Revenue	2989	14110	.	.

Note: This table reports the market share distributions based on cumulative orders at the top for the online and offline export markets. For the online market, we use the census data collected in 2017 (see Section 2.2 for more details on the census data). For the offline market, we use the Chinese Customs data in 2013 for the corresponding HS category of t-shirts (code 6109). We calculate offline firm sizes as their revenue from selling t-shirts. Panel A shows the numbers and market revenue shares of the top listings (products) in children's t-shirts category on AliExpress. The revenue is calculated by multiplying total cumulative orders with current price after discount. Panel B presents the firm-level statistics and compares the online and offline markets. The units for the total revenue is 1000 RMB for the online sales and in million USD for the offline sales.

Table A.4: The Dependence of New Orders on Current Cumulative Orders

Dummy=1 if having an order in the following week	(1)	(2)
Log Orders	0.023***	0.023***
	(0.001)	(0.001)
Constant	-0.009***	-0.010***
	(0.001)	(0.002)
Observations	19855	19680
Store FE	No	Yes

Note: This table reports the results from regressing a dummy variable that equals one for listings that receive orders in the following week on the log number of cumulative orders in the current week. We use the census data at the listing level collected in 2018. See Section 2.2 for more details on the census data.

Table A.5: Summary Statistics of the Children's T-shirts Market on AliExpress: the Baseline Sample

	Observations	Mean	Std Dev	Median	5th Pctile	95th Pctile
			Panel A. Listing Level			
Price	10089	6.14	8.46	5	2.78	11.59
Orders	10089	31.07	189.19	2	0	110
Revenue	10089	163.7	891.68	9	0	636.4
Total Feedback	10089	19.69	127	1	0	67
Rating	5050	96.66	7.4	100	82.9	100
Free Shipping Indicator	10089	.54	.5	1	0	1
Shipping Cost to US	10089	.63	1.44	0	0	2.18
			Panel B. Store Level			
Num of Variety	610	1092.68	1054.38	596	101	3064
Total Orders	610	9287.93	18478.68	3653.5	51	30894
Total Revenue (Discounted)	610	42714.57	82530.27	17946.47	445.56	152587.1
Total Revenue (Undiscounted)	610	50360.67	101736.1	19353.02	0	185653.1
T-shirts Orders	610	286.46	711.96	47	0	1336
T-shirts Revenue (Discounted)	610	1408.14	3932.4	220.27	0	6583.07
T-shirts Zero Sales Indicator	610	.12	.32	0	0	1
Age	610	1.27	1.68	0	0	5
Total Feedback	604	.97	.02	.98	.94	1
Rating	583	4.72	.13	4.7	4.5	4.9
Perc of Positive Feedback	604	.97	.02	.98	.94	1
Communication	583	4.73	.14	4.8	4.5	4.9
Shipping Speed	583	4.64	.17	4.7	4.3	4.8
Positive Feedback in 1 Month	602	806.92	1335.18	389.5	11	2588
Positive Feedback in 3 Months	603	2325.82	3888.12	1063	15	8161
Positive Feedback in 6 Months	604	3857.98	6240.83	1826	16	12959
Neutral Feedback in 1 Month	530	21.37	29.45	12	1	74
Neutral Feedback in 3 Months	550	59.26	85.42	32	2	205
Neutral Feedback in 6 Months	557	102.83	143.85	55	2	368
Negative Feedback in 1 Month	521	22.59	30.44	13	1	79
Negative Feedback in 3 Months	543	59.49	81.81	34	1	208
Negative Feedback in 6 Months	552	104.11	139.34	54.5	2	389
% Positive Feedback in 1 Month	602	.97	.03	.98	.93	1
% Positive Feedback in 3 Months	603	.98	.02	.98	.94	1
% Positive Feedback in 6 Months	604	.97	.02	.98	.94	1

Note: This table reports the same summary statistics as in Table 2 but using the baseline sample collected in May 2018 (see Section 4) for details.

Table A.6: Balance Check

	(1) Control mean/(sd)	(2) T1 mean/(sd)	(3) T2 mean/(sd)	(4) T1-Control b/(se)	(5) T2-Control b/(se)	(6) T2-T1 b/(se)	(7) Joint Test F/(p)
Price After Discount	5.95	5.46	5.63	-0.48	-0.32	0.16	1.28
	(4.10)	(2.57)	(3.72)	(0.29)	(0.35)	(0.29)	(0.26)
Cumulative Orders	0.90	0.73	0.81	-0.18*	-0.09	0.09	0.91
	(1.26)	(1.18)	(1.20)	(0.10)	(0.11)	(0.11)	(0.34)
Total Feedback	0.46	0.38	0.65	-0.08	0.19	0.27*	1.88
	(1.21)	(1.37)	(1.88)	(0.11)	(0.13)	(0.15)	(0.17)
Positive Rating Rate	0.95	0.96	0.91	0.01	-0.04	-0.05	0.83
	(0.21)	(0.17)	(0.28)	(0.04)	(0.04)	(0.05)	(0.36)
Free Shipping Dummy	0.50	0.45	0.48	-0.05	-0.02	0.03	0.21
	(0.50)	(0.50)	(0.50)	(0.04)	(0.04)	(0.05)	(0.65)
Shipping Price	0.74	0.76	0.69	0.02	-0.05	-0.06	0.25
	(1.01)	(0.89)	(1.03)	(0.08)	(0.09)	(0.09)	(0.61)

Note: This table checks whether the order and review treatments are correlated with listing characteristics collected prior to the treatment. See Section 4 for discussion of the experiment sample. The first three columns report the mean and standard deviation of the variables for each treatment group. Columns (4)-(6) show the difference between the two groups and the standard errors of the difference. The column heading indicates which groups are being compared. The last column tests whether the three treatment groups have the same mean. ∗ ∗ ∗: p ¡ 0.01; ∗∗: p ¡ 0.05; ∗: p ¡ 0.1.

Table A.7: Treatment Effects of Order and Review: Without Baseline Controls

	All Destinations		English-speaking		United States	
Order	0.023	0.026*	0.014**	0.015***	0.016***	0.017***
	(0.020)	(0.015)	(0.005)	(0.005)	(0.003)	(0.003)
ReviewXPostReview	0.005	-0.014	0.021	0.016	0.017	0.014
	(0.024)	(0.027)	(0.019)	(0.018)	(0.017)	(0.016)
Observations	10270	10270	10270	10270	10270	10270
Group FE	No	Yes	No	Yes	No	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Controls	No	No	No	No	No	No
Clustered SE at listing level	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the treatment effects of the experimentally generated orders and reviews. The dependent variable is the endline number of cumulative orders, calculated using the transaction data collected in August 2018. The independent variable is the order treatment dummy that equals one for all products in the treatment groups (T1 and T2) and zero for the control group. Column 1 reports the average treatment effect, and Columns 2 to 6 report the quantile treatment effects. Standard errors are in the parentheses. *** indicates significance at 0.01 level, ** 0.5, * 0.1.

Table A.8: Treatment Effects on Ranking

	EnterFirst100Pages		EnterFirst15Pages	
OrderXMonth1	0.012	0.013*	0.004*	0.003*
	(0.008)	(0.007)	(0.002)	(0.002)
OrderXMonth2	0.002	0.003	0.002	0.002
	(0.006)	(0.005)	(0.002)	(0.001)
OrderXMonth3	-0.003	-0.003	-0.000	-0.000
	(0.007)	(0.006)	(0.002)	(0.002)
OrderXMonth4	-0.003	-0.002	0.002	0.002
	(0.009)	(0.008)	(0.002)	(0.002)
Observations	10270	10270	10270	10270
Group FE	No	Yes	No	Yes
Week FE	Yes	Yes	Yes	Yes
Baseline Controls	Yes	Yes	Yes	Yes
Clustered SE at listing level	Yes	Yes	Yes	Yes

Note: This table reports the treatment effects of the experimentally generated orders and reviews on listing ranks using the 13-week panel of the experiment sample. The dependent variable in column 1-2 (3-4) is a dummy variable that equals one if the listing enters the first 100 (15) pages in the no-group search. The baseline controls include the baseline total number of cumulative orders of the store and of the particular product listing. "Order" is a dummy variable that equals one for all products in the treatment groups (T1 and T2) and zero for the control group. "Review" is a dummy that equals one for all products in T2, where we place one order and leave a review on shipping and product quality. "PostReview" is a dummy that equals one for the weeks after the reviews were given (i.e., from week 7 onward). "MonthX" is a dummy variable that equals one for the X-th month after treatment. Standard errors are in the parentheses. *** indicates significance at 0.01 level, ** 0.5, * 0.1.

Table A.9: Heterogeneous Treatment Effect: Quality

	(1)	(2)	(3)	(4)
Order	0.301	0.327*	0.479	0.797
	(0.236)	(0.180)	(0.392)	(0.635)
OrderXServiceQualityIndex	-0.157	-0.075		
	(0.228)	(0.203)		
ServiceQualityIndex	0.290**	0.093		
	(0.119)	(0.125)		
OrderXStdStar			-0.041	0.132
			(0.160)	(0.329)
StdStarRating			0.172**	0.050
			(0.068)	(0.179)
Constant	0.978***	0.889***	0.759	0.002
	(0.320)	(0.230)	(0.542)	(0.758)
Observations	784	784	168	168
Baseline Controls	Yes	Yes	Yes	Yes
Group FE	No	Yes	No	Yes
Clustered SE at listing level	Yes	Yes	Yes	Yes

Note: This table reports the heterogeneous treatment effects of the experimentally generated orders based on quality measures. The dependent variable is the total number of orders net of our own, calculated using the transaction data collected in August 2018. The baseline controls include the baseline total number of cumulative orders of the store and of the particular product listing. "Order" is a dummy variable that equals one for all products in the treatment groups (T1 and T2) and zero for the control group. The standardized quality measures are constructed by standardizing individual quality metrics first and taking their average within each quality type. See 2.2 for details about the quality metrics. Standard errors are in the parentheses. *** indicates significance at 0.01 level, ** 0.5, * 0.1.

Table A.10: Seller Actions After Treatment

	Panel A: Price							
	AdjustPrice		CutPrice		RaisePrice		ΔLogPrice	
Treatment	0.009	0.008	0.013	0.013	-0.004	-0.005	-0.000	-0.001
	(0.018)	(0.016)	(0.011)	(0.012)	(0.015)	(0.012)	(0.007)	(0.006)
Constant	0.055***	0.056***	0.015	0.015	0.040***	0.041***	0.009	0.009*
	(0.014)	(0.013)	(0.009)	(0.009)	(0.012)	(0.010)	(0.006)	(0.005)
Observations	711	711	711	711	711	711	711	711
Group FE	No	Yes	No	Yes	No	Yes	No	Yes

	Panel B: Shipping Cost							
	AdjustShippingCost		CutShippingCost		RaiseShippingCost		ΔLogShippingCost	
Treatment	-0.016	-0.019	0.004	-0.001	-0.020	-0.020	-0.020	-0.019
	(0.032)	(0.028)	(0.026)	(0.024)	(0.023)	(0.022)	(0.036)	(0.037)
Constant	0.264***	0.266***	0.137***	0.140***	0.116***	0.116***	0.016	0.015
	(0.026)	(0.022)	(0.020)	(0.019)	(0.018)	(0.017)	(0.029)	(0.029)
Observations	768	768	768	768	768	768	768	768
Group FE	No	Yes	No	Yes	No	Yes	No	Yes

	Panel C: Product Description and Introduction of New Listings							
	ChangeTitle		ChangeDescription		HaveNewListings		LogNewListings	
Treatment	0.001	0.000	-0.007	-0.010	-0.005	-0.004	-0.092	-0.092
	(0.011)	(0.011)	(0.019)	(0.019)	(0.013)	(0.012)	(0.080)	(0.067)
Constant	0.020**	0.021**	0.078***	0.080***	0.973***	0.972***	3.043***	3.043***
	(0.008)	(0.009)	(0.015)	(0.015)	(0.010)	(0.009)	(0.063)	(0.052)
Observations	769	769	769	769	764	764	764	764
Group FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: This table presents regression results on sellers' responses after treatment using the experiment sample. AdjustPrice is a dummy that equals one for listings that have adjusted their prices within 13 weeks after treatment. CutPrice, RaisePrice, AdjustShippingCost, CutShipingCost, RaiseShippingCost are dummy variables defined in a similar way. ChangeTitle is a dummy that equals one for listings that have updated their product titles within the 13 weeks after treatment. ChangeDescription is a dummy that equals one for listings that have updated their product descriptions within the 13 weeks after treatment; and a set of descriptions include website pictures, pattern type, material, fit, gender, sleeve length, collar, clothing length, item type, color, etc. HaveNewListings is dummy that equals one for a listing if the store to which it belongs has introduced new listings within 13 weeks after treatment; and LogNewListings is the log number of those new listings. Standard errors are in the parentheses. *** indicates significance at 0.01 level, ** 0.5, * 0.1.

B Details on Data and the Experiment

B.1 Measuring Service, Shipping and Product Quality

In order to examine the relationship between quality and growth dynamics in the presence of search and information frictions, we collect a rich set of quality measures on service, shipping and product via (i) direct communication with the sellers and (ii) actual purchases of the products. For each t-shirt variety (of the same design), our quality grading is conducted on all small listings, all medium-size listings with sales between 6 and 50, and the superstar listing (with the largest sales quantity of the variety).

Service Quality. First, we visited the homepage of each store and sent the following message via the platform to engage in pre-transaction service (i.e., inquiry about a particular product):

"Hi, I am wondering if you could help me choose a size that fits my kid, who is 5 years old, 45lbs and about 4 feet. I would also like to know a bit more about the quality of the t-shirt. Are the colors as shown in the picture? Will it fade after washing? What is the material content by the way? Does it contain 100% cotton? The order is a little urgent; how soon can you send the good? Would it be possible to expedite the shipping and how much would that cost? Thanks in advance!"

We then constructed a measure of service quality based on whether the message was replied to within two days, which was true for 69% of the listings.

Shipping Quality. To capture the quality of shipping, we recorded the date of purchase, the date of shipment, the date of delivery, carrier name, and the condition of package. The numbers of days between the date of purchase and the date of shipping, the number of days between the date of shipping and the date of delivery, whether the package was delivered successfully, and whether the package was broken upon delivery are used as alternative measures of shipping quality. The medium numbers of days between purchase and shipping and between shipping and delivery are 3 and 12, respectively. Again, there are considerable variations, especially in sellers' turnaround to ship the products.

Product Quality. We worked with a large local consignment store of children's clothing in North Carolina to inspect and grade the quality of each t-shirt. The owner has over 30 years of experience in the clothing retail business and was invited to grade the quality of the t-shirts.

Each t-shirt was given an anonymous identification number and the owner was asked to grade the t-shirt on 9 quality dimensions, following standard grading criteria used in the textile and garment industry as shown in Panel A of Figure 2. These dimensions include obvious quality defect, fabric durability, fabric softness, wrinkle test, seams (straightness and neatness), outside stray threads, inside loose stitches, pattern smoothness, and trendiness. Most of these metrics, except trendiness, capture differences across t-shirts that are vertical in nature. For example, at equal prices, consumers would prefer T-shirts with more durable fabric, straight seams, and no loose stray threats. The quality examiner grades each t-shirt along the first dimension based on a 0 or 1 scale, and along the other eight dimensions based on a 1 to 5 scale, with higher numbers denoting higher quality. The identification system ensured that the examiner had no information on the purchase price, popularity, and retailer of the t-shirts and whether the t-shirts belonged to our treatment or control group.

In addition, the examiner was asked to price each t-shirt based on her willingness to pay and willingness to sell, respectively. These two additional metrics would reflect not only product quality but also local consumer preferences assessed based on the examiner's retail experience.

T-shirts within the same variety were grouped together for assessment to make sure the grading could better capture within-variety variations. The examiner also conducted two rounds of evaluation that took place several weeks apart to ensure consistency in grading. Panel B of Figure 2 shows examples of the grading and variations across different quality dimensions.

The mean scores vary from 2.6 to 4.2 across different quality metrics. On average, t-shirts scored the worst on inside stray threads and the best on straight seams. Dimensions that record the greatest variations in scores are outside and inside stray threads and pattern smoothness within t-shirt varieties, and pattern smoothness, trendiness, and outside stray threads across varieties.

B.2 The Review Treatments

In our randomized experiment, we group small listings into three groups of different order and review treatments: a control group C without any order and review treatment, T1 which receives 1 order randomly generated by the research team and a star rating, T2 which receives 1 order and 1 detailed review on shipping and product quality in addition to the star rating.

To generate the content of the product and shipping reviews, we use the Latent Dirichlet Allocation topic model in natural language processing to analyze past reviews and construct the messages based on the identified key words. Specifically, the following reviews were provided (randomly) to listings in T2:

Product Quality:

- "Great shirt! Soft, dense material, quality is good; color matching the picture exactly, and I am happy with the design; no problem after washing. My kid really likes it. Thank you!"
- "Well-made shirt. It was true to size. The material was very soft and smooth. My kid really likes the design. I am overall satisfied with it."
- "This shirt is nice and as seen in the photo. It fits my kid pretty well. The material is quite sturdy and colorfast after washing."

Shipping Quality:

- "The shipping was pretty good. Package arrived within the estimated amount of time and appeared intact on my porch."
- "I am pleased with the shipping. It was fast and easily trackable online. The delivery was right on time and the package appeared without any scratches."
- "Fast delivery and convenient pickup, everything is smooth, shirt came in a neat package, not wrinkled. Thank you!"

We leave positive reviews to all listings unless there are obvious quality defect or shipping problems, in which case no review is provided. Of all the orders placed, about 8 percent have obvious quality defect or shipping issue.

C Baseline Model

We focus on a single product segment (e.g., children's t-shirts). On the supplier side, we assume that the platform features N single-product sellers in the market, without entry or exit. Seller i is endowed with an initial visibility $v(s_i^0)$, where cumulative orders s_i^0 is assumed to be zero for all sellers in our baseline setup.[30]

$$\mathbf{v}^0 = (v^0, ..., v^0),$$

we restrict the function $v(.)$ such that $v^0 \in \mathbb{R}_{++}$ for all $i \leq N$. We denote the cumulative orders vector at the beginning of period t by $\mathbf{s}^t = (s_1^t, ..., s_N^t)$.

On the consumer side, we assume that there is one consumer arriving at the market each period t and purchase from one seller. The probability that the consumer buys from seller i depends on seller i's visibility relative to those of other sellers at the beginning of period t. Specifically,

$$P\left(\mathbf{s}^{t+1} = \mathbf{s}^t + \boldsymbol{e}_n^t\right) = \frac{v_i^t}{\sum_{j=1}^N v_j^t}.$$

We follow Drinea et al. (2002) and Mitzenmacher, Oliveira, and Spencer (2004) to assume that the visibility takes the functional form

$$v(s_i^t) = (s_0 + s_i^t)^\zeta$$

where s_0 can be an arbitrarily small positive number, and ζ governs the sensitivity of each seller's visibility to its cumulative orders.

Given each seller's cumulative orders, it is straightforward to define their corresponding market share vector at the beginning of each period t as

$$\mathbf{m}^t = (m_1^t, ..., m_N^t),$$

where $m_i^t = \frac{s_i^t}{\sum_{i=1}^N s_i^t}$. Drinea et al. (2002) and Mitzenmacher, Oliveira, and Spencer (2004) show that ζ impacts the limiting distribution of seller's market share. First, when $\zeta < 1$, the probability of a seller making a sales increases less than proportionally with the seller's cumulative orders and the seller's visibility advantage dissipates over time. As a result, the limiting market share distribution tends to be uniform.

Second, when $\zeta > 1$, the probability of a seller making a sales increases more than proportionally with its cumulative orders. This amplifies its initial advantage in visibility over time. Thus, one seller will eventually obtain monopoly; that is, there exists a time after which all subsequent consumers buy from just one seller.

Finally, when $\zeta = 1$, the process of cumulative orders $\left(\mathbf{s}^t\right)_{t=1,2,...}$ follows a Polya urn process. When t goes to infinity, the limiting distribution of market shares is a Dirichlet distribution where the parameters

[30] More generally, we can allow for any positive number of s_i^0 to accommodate potential spillovers from a seller's other previous platform activities.

only depend on s_0. Specifically,

$$\mathbf{m}^t \xrightarrow{d} \text{Dir. } (s_0, ..., s_0), \text{ as } t \to \infty.$$

The above comparison suggests that the case $\zeta = 1$ is the most aligned with the market share distribution observed in our sample. This is further confirmed by the experimental evidence which suggests that initial demand shocks can play a significant role in sellers' future performance.

Define the vector of order statistics of market shares $\mathbf{m}^t$ as

$$\tilde{\mathbf{m}}^t \equiv \left(m^t_{(1)}, \ldots, m^t_{(N)} \right),$$

where $m^t_{(i)}$ is the i-th largest value among all the sellers' market shares in period t. Assuming that the initial visibility s_0 is small and that the total number of sellers N is fairly large, $\tilde{\mathbf{m}}^t$ can be showed to converge to a Poisson-Dirichlet distribution that is completely determined by $\lambda \equiv N \cdot s_0$. Intuitively, this result suggests that only the relative magnitude of N and s_0 matters for the limiting market share distribution. Therefore, we fix the number of sellers at $N = 1000$ and estimate s_0 as a model parameter.

D Details on the Simulation Procedure

We implement the Method of Simulated Moments according to the following procedure.

D.1 Recover Marginal Cost

In the first step, We use the data distributions of price, review, and cumulative orders to recover the distribution of costs, F_c, relying on the set of first order conditions from the sellers' static pricing problem that is described in section 5.2. We simulate demand $D_i(\mathbf{p}, \mathbf{r}, \mathbf{s})$ and demand derivative $\frac{\partial D_i}{\partial p_i}(\mathbf{p}, \mathbf{r}, \mathbf{s})$ based on equation (3) and (5).

D.2 Initialize Sellers in the Market

We initialize the market by setting the cumulative orders of sellers at 0 and the visibility of sellers at $v_0 = s_0 > 0$. In addition to the marginal distribution of costs F_C obtained in step D.1 and the standard normal marginal distribution of quality, we use the Gaussian Copula to model their dependence. Specifically, we draw the tuple (q, c) for each seller according to the following steps:

1. Draw a vector $\mathbf{Z}$ from the multivariate standard normal distribution with correlation ρ,

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right).$$

2. Calculate the standard normal CDF of $\mathbf{Z}$:

$$U_1 = \Phi(Z_1), \quad U_2 = \Phi(Z_2).$$

3. Transform the CDF to quality and cost values using their marginal distributions:

$$c_{\text{draw}} = F_C^{-1}(U_1), \quad \mathrm{q}_{\text{draw}} = \Phi^{-1}(U_2) = \Phi^{-1}(\Phi(Z_2)) = Z_2.$$

After drawing the cost and quality for each seller, we solve their static pricing problem to set the initial prices.

D.3 Simulate One Period

In each period, we use the weighted sampling without replacement to generate the consumer's search sample of size K and draw idiosyncratic preference ε from the I.I.D. type I extreme value distribution. Based on the average reviews, we calculate the expected quality and the expected utility of purchasing from each seller in the search sample. Then, we simulate the purchasing decision, the realized experience for the consumer, and the review he/she leaves. At the end of each period, we update the cumulative orders and the average review for the seller that has made a new sales. In addition, we allow the sellers to update their prices by solving the static pricing problem at the frequency that matches with the observed frequency of price adjustment.

D.4 Simulate Moments

Starting from the initialized market, we repeat step D.3 for $T = 10000$ times so that the market share based on cumulative orders reaches an invariant distribution. Then, we simulate forward for another ΔT periods to produce moments from a stabilized market. Specifically, we calculate the distribution of cumulative revenue for the sellers, the regression coefficient of log price and quality, and the regression coefficient of the market share of cumulative orders on quality in the final period, i.e. $t = T + \Delta T$. And we calculate the dependence of seller's new order on cumulative orders using simulated data from period $T + 1$ to $T + \Delta T$.

D.5 Weighting Matrix and Objective Function

We bootstrap our data sample moments 1000 times and construct the weighting matrix W. The objective function used for optimization is

$$Q(\theta) = -\frac{1}{2}\left(g_0 - \gamma_m(\theta)\right)' W \left(g_0 - \gamma_m(\theta)\right),$$

where g_0 is the data moments vector, $\gamma_m(\theta)$ is the simulated moments vector based on $m = 100$ simulations, and $\theta = (s_0, \sigma, \rho, K, \gamma)$ is the vector of parameters.