E-Commerce Integration and Economic Development: Evidence from China*

Victor Couture[†], Benjamin Faber[‡], Yizhen Gu[§] and Lizhi Liu[¶]

February 2018

Abstract

The number of people buying and selling products online in China has grown from practically zero in 2000 to more than 400 million by 2015. Most of this growth has occurred in cities. In this context, the Chinese government recently announced the expansion of e-commerce to the countryside as a policy priority with the objective to close the rural-urban economic divide. As part of this agenda, the government entered a partnership with a large Chinese e-commerce firm. The program invests in the necessary logistics to ship products to and sell products from tens of thousands of villages that were largely unconnected to e-commerce. The firm also installs an e-commerce terminal at a central village location, where a terminal manager assists households in buying and selling products through the firm's e-commerce platform. This paper combines a new collection of survey and administrative microdata with a randomized control trial (RCT) that we implement across villages in collaboration with the e-commerce firm. We use this empirical setting to provide evidence on the potential of e-commerce integration to foster economic development in the countryside, the underlying channels and the distribution of the gains from e-commerce across households and villages.

Keywords: E-commerce, trade integration, economic development, rural-urban divide
JEL Classification: F63, O12, R13

*We are grateful to CEGA, the Clausen Center, the Haas School of Business, the Weiss Family Fund and the Bill and Melinda Gates Foundation for providing funding for this study. We have benefited from outstanding research assistance by Hero Ashman, Wenwei Peng, Jose Vasquez-Carvajal and Yi Wei. We are grateful to Lijun Sun, Wei Wang, Wei Zheng and Fang Ye for their tireless support during the implementation of this project. We also thank Hongbin Gao, Zhenzhong Sheng, Liang Chen, Wentao Zhang and Zhengwei Jiang for providing us access to and assistance with the transaction database. We are grateful to David Atkin, Jie Bai, Loren Brandt, Lauren Falcao-Bergquist, Lorenzo Casaburi, Dave Donaldson, Pablo Fajgelbaum, Thibault Fally, Frederico Finan, Cecile Gaubert, Edward Glaeser, Paul Gertler, Jonas Hjort, Amit Khandelwal, Supreet Kaur, Michael Kremer, David Levine, Ted Miguel, Mushfiq Mobarak, Ferdinando Monte, Ben Olken, Ralph Ossa, Andres Rodriguez-Clare, Daniel Sturm, Matt Turner, Eric Verhoogen, Jonathan Vogel, Michael Walker, Shangjin Wei, Christopher Woodruff, Daniel Xu, Noam Yuchtman, Fabrizio Zilibotti and participants at several seminars and conferences for very helpful comments. All views expressed and errors are our own.

[†]Real Estate Group, Haas School of Business, UC Berkeley
[‡]Department of Economics, UC Berkeley and NBER
[§]Institute for Economic and Social Research, Jinan University
[¶]Department of Political Science, Stanford University

1 Introduction

The number of people buying and selling products online in China has grown from practically zero in the year 2000 to more than 400 million by 2015, surpassing the US as the largest e-commerce market in terms of users and total sales.[1] Outside of China, a growing number of developing countries, especially in Asia, Eastern Europe, Latin America and the Middle East, are experiencing rapid growth in e-commerce activity (WTO, 2013; UNCTAD, 2016b). Most of this growth to date has taken place in the cities of the developing world. In this context, the Chinese government recently announced the expansion of e-commerce to the countryside as a national policy priority to foster rural economic development and reduce the rural-urban economic divide.[2] Other developing countries, such as Egypt, India and Vietnam, have recently announced similar policies to invest in the expansion of e-commerce trading into rural areas, where the majority of the population live.[3]

To date, these policies have been motivated mainly by a number of prominent case studies of highly successful "e-commerce villages" that have experienced rapid output growth by selling both agricultural and non-agricultural products to urban markets via e-commerce. For example, by the end of 2017, China's largest e-commerce platform, Taobao, had branded more than 2000 rural locations in China as so called "Taobao villages", based on their concentration of online sellers and sales volumes on the firm's platform (AliResearch, 2017).[4] Inspired by these success stories, much of the current policy focus has been on rural producers. By lowering trade and information costs to urban markets, the arrival of e-commerce is meant to increase rural incomes through higher demand for local production and stronger incentives for rural entrepreneurship. There has been much less emphasis on the potential benefits to rural consumers. However, recent descriptive evidence from urban China suggests that e-commerce demand is strongest in smaller and more remote cities, pointing to potentially significant consumer gains in rural areas.[5]

Despite the fast growth of e-commerce, we currently have limited empirical evidence on the economic consequences of access to e-commerce trading in developing countries. The recent growth of a number of "e-commerce villages" in China has captured the imagination of policy-makers and the general public, but important questions remain about whether market integration through new online trading platforms can have a broad and significant impact on rural develop-

[1]See e.g. PFSweb (2016) and Statista (2016).

[2]The so-called "Number One Central Document" sets out yearly strategic priorities of the Chinese central government (the Central Committee of the Communist Party of China and the State Council in particular). The expansion of e-commerce to the countryside has featured in this document each January since 2014.

[3]As part of "Digital India", a collaboration between the Ministry of Electronics and IT and India Post have been tasked to expand online buying and selling in rural India (MEITY, 2016). Other recent examples include Egypt's National E-Commerce Strategy (MCIT, 2016) and Vietnam's E-Commerce Development Masterplan (PM, 2016). Following this policy interest, UNCTAD recently announced the launch of a new policy platform, "eTrade For All: Unlocking the Potential of E-Commerce in Developing Countries", to provide technical assistance and funding for e-commerce expansions in the developing world (UNCTAD, 2016a).

[4]E-commerce villages have also received much press coverage. See e.g. "China's Number One E-Commerce Village" (BBC Global Business, 01 May 2013), "Inside China's Tech Villages" (The Telegraph, 05 Nov 2016), "Once Poverty-Stricken, China's "Taobao Villages" Have Found a Lifeline Making Trinkets for the Internet" (QZ, 01 Feb 2017), "Chinese 'Taobao Villages' Are Turning Poor Communities into Huge Online Retail Hubs" (Business Insider, 27 Feb 2017).

[5]In the US, the share of e-commerce in total US retail sales is estimated to be about 10-15 percent (e.g. FRED (2016)). In China, McKinsey (2016) reports this share to be as high as 20-30 percent in smaller cities, and Fan et al. (2016) find that this share increases by on average 1.2 percent as city population decreases by 10 percent.

ment. To answer these questions, this paper combines a randomized control trial (RCT) that we implement across villages in collaboration with a large Chinese e-commerce firm with a new collection of household and store price survey microdata and the universe of transaction records from the firm's internal database. We use this empirical setting to provide evidence on the potential of e-commerce integration to foster economic development in the countryside, the underlying economic channels, and the distribution of the gains from e-commerce across households and villages. These findings can serve as a first step towards building a rigorous evidence base on the economic consequences of rapid e-commerce growth in developing countries.

E-commerce is the ability to buy and sell products through online transactions coupled with transport logistics for local parcel delivery and pickup from the producer. Bringing e-commerce to the countryside in developing countries requires more than internet access. As in many emerging countries, the internet has spread rapidly to most parts of the Chinese countryside due to both smartphones and expanding broadband access. Instead, there are two currently binding barriers to e-commerce trading in the Chinese countryside, which we refer to as the logistical and the transactional barriers. First, the logisical barrier relates to the the lack of modern commercial parcel delivery services. These distribution networks have started operating in Chinese cities, but have not started servicing large parts of the countryside. One well-known challenge to rural delivery is the so called "last mile" between urban logistical hubs and relatively small pockets of population in the countryside. Second, many rural residents potentially face a transactional barrier, due to lack offamiliarity with navigating online platforms and access to online payment methods. In addition, villagers may not trust transactions that occur before inspecting the product or without interacting with buyers in person.

To overcome these barriers, the Chinese government recently partnered with a large firm that operates a popular Chinese e-commerce platform. The program aims to invest in the necessary transport logistics to offer e-commerce in rural villages at the same price, convenience and service quality that buyers and producers face in their county's main city center. To this end, the e-commerce firm builds warehouses as logistical nodes for rural parcel delivery near the urban center, and fully subsidizes transport between the county's city center to and from the participating villages. To address the transactional barrier, the program installs an e-commerce terminal in a central village location. A terminal manager, who is employed by the firm, assists villagers in buying and selling products through the firm's e-commerce platform, and villagers can pay upon receipt of their products or get paid upon pickup of their shipments in cash at the terminal location. From the end of 2014 to June 2016, approximately 16,500 Chinese villages in 333 counties and 27 provinces had been connected to e-commerce through the program. This expansion continues at the time of this writing, with an internal goal of 40,000 villages in 600 counties by the beginning of 2018.

Theoretically, we rationalize the program as a reduction in trade and information costs between participating villages and the rest of urban China that is already connected to e-commerce. An advantage of this setting is that we can study the reduction in trade frictions through e-commerce without confounding the counterfactual with the consequences of first-time internet access more broadly. The participating villages were already connected to the internet, and the program makes no changes on this front. Furthermore, the program only directly affects trading partners through e-commerce, while other trade costs, e.g. to control villages, remain unchanged. The RCT and data analysis that we describe below exploit this empirical setting to provide evidence on the local economic effects of e-commerce trading access on rural households.[6]

Our analysis proceeds in four steps. In the first step, we derive a general expression to quantify the program's effect on household economic welfare, that guides the survey data collection and empirical analysis. Since the treatment we are interested in evaluating can affect not just individual behavior and the nominal earnings of households, that we can in principle record directly as part of the survey data collection, but also local household price indices in the denominator of real incomes, the evaluation of the welfare impact requires theoretical structure on the demand side. In particular, some of the potential effects on household cost of living are likely to occur at the extensive margin of consumer choice, such as the arrival of a new e-commerce shopping option or local store exit. For such changes in the availability of local consumer options, the effective price changes are unobserved since no information exists at either baseline (new options) or endline (disappearing options) survey periods. Following a revealed-preferences approach in industrial organization (e.g. Hausman (1996)) and international trade (e.g. Feenstra (1994); Atkin et al. (in press)), we make use of observed substitution of household spending into new options or away from disappearing ones to infer the effective change in consumer welfare across different product groups. More generally, the welfare expression allows us to break down the overall effect of e-commerce integration into several distinct components that we can link to the microdata. We also discuss the assumptions under which rural-to-rural general equilibrium (GE) spillover effects are negligible. In the analysis, we begin by estimating simple differences in outcomes between treatment and control villages under this baseline assumption, and then use two types of additional empirical moments to investigate the presence of spillovers across villages.

In the second and third steps, we estimate the empirical moments of this welfare expression using household and village survey microdata, as well as the firm's internal database. The RCT takes place in 8 counties located in three provinces, Anhui, Henan and Guizhou, that have a large share of rural population. For each county, we were given authorization to randomly select control villages from a list of candidates that had been extended by 5 villages per county for the purpose of this research. Upon receipt of this extended list of village candidates, we randomly select 5 control villages and 7-8 treatment villages from each county's list. The remaining villages on the list also receive a program terminal as planned. Our sample thus includes 40 control villages and 60 treatment villages across the 8 counties, which we selected from a total number of 432 candidate villages (on average 54 villages per county). Terminal installation and local e-commerce deliveries and pick-ups proceed shortly after we complete the baseline data collection.

We complement this experimental design with survey data that we collect from households and local retail establishments. We collect baseline data in 8 different counties in December 2015, January 2016 and April-May 2016 for 2800 households (roughly 8600 individuals) in the 100 villages. For the endline, we collect data from the original household sample, and also extend the number of households by 10 randomly selected households (leading to an endline sample of

[6]It would be outside the scope of this study to also attempt a social cost-benefit analysis of the program, which would require additional detailed (and confidential) information on the cost side from both the government actors and the e-commerce firm.

3800 households).[7] For each household, we collect detailed information about consumption expenditures (e-commerce and other), expenditures on production inputs, economic activities and incomes. We also collect baseline and endline information on 115 local retail price quotes for each village at the barcode-equivalent level across 9 consumer product groups as well as for business/production inputs. Overall, these survey data are aimed at quantifying the effect on household real incomes, and to distinguish between a number of underlying channels for both consumption gains (the denominator) and production-side effects (the numerator). In terms of timing, we conduct the endline data collection 12 calendar months after the baseline in each county.[8] This implies that the RCT and survey-based data allow us to quantify the program's effect up to 12 months after the arrival of e-commerce. In order to lift this and some other practical constraints of the fieldwork, and investigate the extent to which the censoring of outcomes one year after the intervention may mask longer-term adjustments on both the consumption and production sides, we are able to combine the fieldwork with additional evidence from the firm's administrative database, to which we turn in the next step.[9]

In the third step, we bring to bear the firm's internal database, that provide us with access to the universe of village transaction records in 5 provinces (including the three above) over the period between November 2015 until April 2017. The database covers roughly 27.8 million purchasing and selling transaction records for about 12,000 village terminals that were in operation over this 18-month period.[10] The data allow us to observe village-level purchases and online sales up to two years and 4 months after the arrival of e-commerce. We use these data to answer four questions that are outside the scope of the RCT and survey data collection: i) to what extent are the RCT sample villages representative of the program villages in the Chinese countryside more broadly?; ii) to what extent are the results from the endline survey data sensitive to monthly seasonality?; iii) what is the time path of adjustment for e-commerce buying and selling each month since program entry, and do the effects increase for periods more than one year post-installation?; and iv) to what extent are the survey data missing very successful, but rare, tail events on the production side that could affect the mean impact on household incomes per capita?

In the final step, we use the empirical estimates from steps 2 and 3 in combination with the theoretical framework in step 1 to quantify the impact of the program on average household economic welfare, the underlying channels and the distribution across households and villages.

[7]In one of the 8 counties, the local government suspended further activity by our teams after we had completed endline data collection for 8 of the 12 sample villages. This was unrelated to our operation, which followed the same protocol as elsewhere. As a result, we have endline data for 96 instead of the 100 villages. As the timing of data collection within the county was random, the 4 missing villages are not particular in any way. They included 2 control villages and 2 treatment villages.

[8]The fast pace of the program's expansion places bounds on the timing of the endline. After the baseline data collection, additional waves of implementation started appearing on the county teams' schedules. Given that our control villages were selected from a list of promising candidate villages, they ranked highly for additional installations that started being rolled out one year after the initial wave.

[9]Related to this, we note that much of the existing literature on the consequences of ICT in developing countries have estimated effects after relatively short periods of time. For example, Jensen (2007) documents significant effects of Indian cell phone towers on local market prices and other outcomes within weeks of installation. More recently, Hjort & Poulsen (2017) use quarterly and annual data from several African countries and document effects of fast-speed internet on local employment and incomes that arise within 3-12 months post-installation. Also see related literature further below.

[10]As we discuss below, the out-shipment data cover 16 months starting in January 2016, rather than November 2015 as for the purchase transactions.

We find that the program leads to sizable gains in real incomes among the group of rural households who are induced to use the e-commerce terminal. These users represent about 14 percent of the rural household sample and 13 percent of the village population after adjusting for sampling weights. For the average rural household, including non-users, these gains are statistically significant but more muted. Underlying these effects, we find strong heterogeneity across households and villages. Beneficiaries are on average significantly younger, richer, live in closer proximity to the e-commerce terminal, and in villages that are relatively more remote. Conditional on these characteristics, we do not find evidence that household education or the test scores of the terminal managers affect the extent of household gains from e-commerce.

In terms of channels, we find significantly stronger economic gains among villages that were not previously serviced by commercial parcel delivery, suggesting that the program's gains are mainly due to overcoming the logistical barrier, rather than the transactional one. On the consumption side, we find that the e-commerce terminals offer on average lower prices, higher convenience and increased product variety compared to pre-existing local retail choices, both within the village and in nearby towns. The gains in household purchasing power are strongest for durable product groups, such as electronics and appliances. We also find suggestive evidence that the program led to additional product variety in pre-existing local stores, as their managers source new products through e-commerce. On the other hand, we find no evidence of significant pro-competitive effects on local retailer prices. On the production side, we find no evidence for significant effects on the local economy: online selling activity, purchases of production inputs, household incomes and entrepreneurship are not significantly affected by the arrival of the program. Overall, we find that the gains from e-commerce are driven by a reduction in local household cost of living that is mainly due to the direct gains from access to the new e-commerce shopping option for local households. These gains are in the order of a 5 percent reduction in the cost of living for retail consumption among users, and about a 1 percent reduction for the average household living in these villages. For durable good consumption, the estimated reduction in the local cost of living is 17 percent among users and on average 3 percent among all households.

Using the firm's database, we find little evidence on the consumption side suggesting that the adjustment takes longer than one year: the consumption-side uptake materializes within 2-4 months of entry and then remain mostly constant over time. On the production side, we find evidence that village-level out-shipments significantly increase over time beyond the 12-month survey window. However, the effect on total out-shipments remains relatively minor after more than two years post-program entry, with a small upper-bound effect on local household incomes. Related to this, we do not find evidence that the survey data fails to pick up highly successful but rare tail events on the production side that could in principle shift the mean effect on local household outcomes.

Overall, our findings suggest that e-commerce trading access offers significant economic gains to certain groups of the rural population, rather than being broad-based. Compared to the recent case studies highlighting a set of highly successful rural e-commerce production hubs, our analysis reveals that a quite particular mix of local factors must be underlying these prominent success stories. In the absence of complementary interventions, such as for example business training, access to credit or targeted online promotions, we show that large and significant production-side effects appear unlikely to materialize for the average rural market place in the short to medium-run. These findings can serve to inform the current wave of policy interest in e-commerce expansions and research in this area. In particular, a better understanding of the factors and potential complementary interventions that enable some markets to thrive on the production side under e-commerce integration seems a promising agenda for future work in this area.

This paper relates and contributes to the recent literature on globalization and development using within-country empirical variation (e.g. Topalova (2010); Donaldson (in press); Atkin et al. (in press)). Given the empirical context, we also relate to the recent literature on the consequences of transport cost reductions within China (e.g. Banerjee et al. (2012); Baum-Snow et al. (2016); Faber (2014)). Instead of focusing on trade liberalization, transport infrastructure or the effects of FDI, we set focus on the economic consequences of e-commerce, a recent but fast-growing channel of market integration in developing countries that has so far received relatively little attention in the literature.

Our findings also relate to the literature on the consumer gains from e-commerce (e.g. Brynjolfsson et al. (2003); Goldmanis et al. (2010); Einav et al. (2017)), and cost of living as a function of city size and urban density (e.g. Handbury (2013); Handbury & Weinstein (2015); Couture (2016); Fan et al. (2016)). In this literature, we most closely relate to recent work by Fan et al. (2016) who use data on e-commerce sales on the Taobao platform across 315 prefectures in China for the year 2013 to document a decreasing relationship between prefecture population and online expenditure shares in the cross-section. These findings would suggest that the consumer gains from e-commerce are expected to be the largest among small and remote market places. Relative to the existing literature, this paper uses experimental variation in the arrival of e-commerce to the countryside to quantify the effects on both consumption and production. Our findings suggest that the relationship between e-commerce usage and city size does not appear to hold monotonically as we move from relatively large urban centers further to the left tail of the population size distribution in the countryside.

The paper is also related to the recent literature on the effects of the internet in developing countries. Goyal (2010) studies the consequences of the introduction of internet kiosks with wholesale price information in the Indian state of Madhya Pradesh, and finds a positive effect on local soy prices and area under soy cultivation. More recently, Hjort & Poulsen (2017) use the geography of existing terrestrial communication networks in Africa in combination with the timing of submarine internet cable connections to study the effect of fast-speed internet on labor markets in several African countries. They find a positive effect on overall employment that is mainly driven by an expansion of higher-skill employment.[11] Relative to the existing literature, this paper sets focus on a different question of policy interest. Rather than estimating the effects of the internet more broadly, we explore the consequences of the arrival of e-commerce among rural

[11] There is also a large empirical literature using natural experiments to estimate the socioeconomic effects of the internet in the US and Europe. e.g. Bhuller et al. (2013) and Akerman et al. (2015) exploit the timing of the roll-out of broadband internet across Norwegian municipalities to estimate the effect on sex crime and skilled-to-unskilled wages and productivity respectively. Forman et al. (2012) use US county-level measures of infrastructure costs, industry mix and characteristics of nearby locations as instruments for internet investment to estimate the effect on local wages. Campante et al. (2013) interact municipality distances to the broadband backbone with the timing of the national broadband roll-out to estimate the effect on political participation in Italy. Falck et al. (2014) exploit municipality-level variation in distances to telephone exchange stations for a sample of German municipalities that do not host a station to estimate the effect on political participation.

Chinese markets. Since the expansion of e-commerce requires specific investments to overcome both logistical and transactional barriers beyond the provision of internet access, our analysis can serve as a first step to inform the current wave of policy interest in the promise of e-commerce as a driver of rural development.

Finally, our findings relate to recent literature on the sources of the rural-urban economic divide in developing countries (e.g. Young (2013); Lagakos et al. (2016); Hamory et al. (2016)). A central question in this literature is the extent to which features of locations, rather than the selection of people across space, can explain the observed rural-urban gap in economic development. Our findings suggest, perhaps surprisingly, that in the Chinese case a lack of urban market access –a characteristic that differs between rural and urban locations– does not by itself appear to be a strong factor explaining observed disparities between the countryside and urban centers, at least in the short to medium-run. In this respect, our findings complement existing evidence suggesting that selection plays an important role in rationalizing observed differences in rural and urban economic development.

The remainder of the paper is structured as follows. Section 2 describes the context, experimental design and data. Section 3 presents the theoretical framework. Section 4 presents the empirical analysis based on the RCT in combination with the survey data. Section 5 provides additional evidence using the firm's internal database. Section 6 presents the welfare quantification. Section 7 concludes.

2 Context, Experimental Design and Data

2.1 Context and Program Description

Following the announcement of the policy objective to expand e-commerce to the Chinese countryside as part of the so-called Number One Central Document in January 2014, the Chinese government entered a partnership with a large firm that operates a popular Chinese e-commerce platform. The program makes two main types of investments to enable villagers to buy and sell online through the firm's platform. First, the program invests in the local distribution network, which the firms views as a necessary condition to provide e-commerce access in rural areas. Before the arrival of the program, most villages were not serviced by commercial parcel delivery operators who had not solved the problem of the "last mile" transportation between dispersed rural households and urban county centers.[12]

The program sets out to change this lack of service with logistics investments targeted at e-commerce. In particular, the firm oversees the construction of warehouses in the counties that serve as logistical nodes to pool all e-commerce related transportation requests to and from the participating villages. These warehouses are located close to the main urban center of the county with good cross-county transport access. The program also fully subsidizes the transportation cost between these warehouses and the participating villages so that rural households face the same delivery costs and prices as households in the urban parts of the county. The rationale for this subsidy is that village deliveries and pickups start from a low basis, which due to economies of scale in rural transportation makes the starting phase of e-commerce prohibitively costly for

[12]To receive packages via mail in absence of commercial parcel delivery services, rural households have to travel to the county or township center to pick up the package after receiving notification by mail that it has arrived.

village customers despite the investments in warehouses. The calculation of the government and the firm is that as the scale of rural e-commerce grows, per unit transport costs will decline enough to remove the need for a subsidy. Neither the warehouses, nor the last-mile subsidy can be used for shipments outside of the firm's e-commerce platform.

The second investment is the installation of a program terminal in a central village location. The e-commerce terminal is a PC, keyboard and mouse connected to a flat-screen monitor mounted on the wall of a dedicated shop space and displaying the firm's website. On the screen, consumers and producers can choose their purchases or see their sales requests on the platform. The firm employs a terminal manager to assist local households in buying and selling products through the firm's e-commerce platform. The terminal manager receives a reward of about 3-5 percent for each transaction completed through the terminal. Before deciding on terminal installations, the firm solicits applications from potential local store operators and schedules an exam for the applicants. The score of this exam is one of the criteria that the firm uses to determine whether a village is a candidate. Villagers can pay in cash when the products arrive at the store for pickup, or they get paid upon delivery of their products for pickup at the store location if selling online. Instead of using the terminal interface, households can also choose to use the firm's e-commerce platform remotely on smartphones or PCs to order product deliveries or pickups at the terminal location. When referring to terminal usage below, we include all types of use of the e-commerce platform. The firm views the village terminals as overcoming three challenges. First, local households may not be used to or comfortable with navigating online platforms. Second, they often do not have access to online payment methods. And third, they may not trust online purchases or sales before inspecting the goods in person or having interacted with buyers directly.

2.2 Sample, Design and Data

In this section, we briefly describe the sample, experimental design and data used in the analysis. Figure 1 presents a map of the locations where the RCT takes place. Appendix C provides additional details on surveyor training, data quality management, sampling and variable construction. And Tables 1-3 and A.1-A.4 present descriptive statistics.

Selection of Provinces and Counties

There are two main factors determining our survey location in Anhui, Henan and Guizhou, and the 8 counties within these provinces. First, our survey location depended on the timing of the program's roll-out across different provinces and counties, which had been decided before our collaboration with the firm. Second, we were guided by the internal evaluation of the program's senior managers as to whether the provincial and county managers in question would be willing to cooperate with our research protocol. These counties are: Huoqiu (Anhui), Linying (Henan), Linzhou (Henan), Minquan (Henan), Suixi (Anhui), Tianchang (Anhui), Xifeng (Guizhou) and Zhenning (Guizhou). In Section 5, we are also able to investigate the representativeness of our sample villages relative to all participating villages using the firm's internal transaction data in 5 provinces over this period.

Selection of Villages and Experimental Design

The unit of randomization is the village. For each county, we obtain a list of candidates that had been extended by 5 promising village candidates that would have not been part of the list in absence of our research. The two main factors determining the village selection within a county from the firm's operational perspective are i) a sufficient level of local population, ii) accessibility by roads, and iii) the presence of a capable store applicant (as measured by the applicant's test score). Overall, the pool of selected villages for participation in the program, based on which we are able to implement randomization, is not a random sample of the Chinese countryside, but is instead likely a positively selected group of villages with better expected conditions for e-commerce usage in both consumption and production. We return to this when discussing our findings in the conclusion.

Upon receipt of this extended list of village candidates for each county, we randomly select 5 control villages and 7-8 treatment villages. The remaining villages on the extended list receive program terminals as planned. The full sample thus includes 40 control villages and 60 treatment villages across the 8 counties, which we selected from a total number of candidates of 432 villages that we received in the extended listings from the 8 operations teams (on average 54 villages per county).

We restrict the list of villages entering the stratification and randomization to villages with at least 2.5 km distance to the nearest village on the county list, where possible.[13] We then stratify treatment and control villages along four dimensions. First, we balance the selection of treatment and control to both have a ratio of 85:15 with respect to pre-existing availability of commercial package delivery (85% not available, 15% available), which is close to the observed ratio among all candidate villages. We obtain information on the availability of commercial package delivery for each village on the candidate list from the program's local county managers (who is not aware what we require that piece of information for). As we discuss below, having villages in our sample with pre-existing commercial delivery services allows us to further investigate the effect of the program that is driven by the terminal access point (i.e. the effect of lifting only the transactional barrier), relative to the effect of providing both the terminal access point and the necessary logistics for local e-commerce deliveries and pick-ups (i.e. the effect of lifting both the transactional and logistical barrier to e-commerce). We further stratify the selection of treatment and control villages on the basis of the equally-weighted average of the z-scores for three village variables: the local store applicants' test score, the village population, and the ratio of non-agricultural employment over the local population. We obtain the last variable from the establishment-level data of the Chinese Economic Census of 2008 which surveys every non-agricultural establishment in the counties.

Once we obtain the candidate list for each county, we have about 2-3 weeks to run the randomization and send in the survey teams for data collection in 5 control villages and the 7-8 treatment villages. After that, terminal installations take place and e-commerce begins in the treatment villages. Compliance with our assignments to treatment and control villages is not perfect: the program was rolled out in 38 of the 60 villages assigned to treatment, and it was present in 5 out of the 40 control villages.[14] We therefore report both intent-to-treat and treatment-on-treated ef-

[13] In counties with relatively short candidate lists we had to marginally extent this threshold, leading to a small number of villages with less than 2.5 km distances to the nearest other villages on the candidate list. The mean and median distances for villages without terminals to the nearest terminal location were 10.6 and 9.1 km respectively. We return to this discussion as part of the spillover analysis in Section 4.4.

[14] As discussed below, in our estimation sample we were able to collect data for 96 of the 100 villages during the endline survey. The treatment proportions for the sample of 96 villages are 37/58 and 4/38 respectively.

fects. The main reason for imperfect compliance is that we can randomize treatments only at the stage before the terminal manager candidates get to accept the offer and sign the contracts. Not all candidates that apply and make it to the list of viable candidates (from which we randomize) end up accepting the offer and signing the contract after we send the results of our randomization back to the local operation team. As a result, not all of the 60 chosen candidate villages end up with effective program roll-out. On the other side, the small number of control villages that end up with terminals are either due to mis-communication between our research team and the local operations teams or due to local political constraints (e.g. a county government official insisting on a particular village).

Tables 1 and 2 and appendix Tables A.1-A.4 present descriptive statistics of the baseline data at the individual level, the household level and for local retail prices. The experimental design appears to have been successful in creating treatment and control groups that are on average balanced in terms of pre-existing outcomes. As discussed in Section 4, our empirical analysis will also condition on the baseline values of the outcomes to be tested.

Sampling of Households and Household Survey Data

For the first round of data collection (December 2015 and January 2016 in Anhui and Henan, and April and May 2016 in Guizhou), we collect data from 28 households per village. 14 of those households are randomly sampled within a 300 meter radius of the planned terminal location ("inner zone"), and 14 households are randomly sampled from other parts of the village ("outer zone"). The household survey respondent is the member with the most knowledge of household consumption expenditures and income. Each respondent receives a gift to thank them for their participation in the survey (e.g., box of premium sweets, soaps, hand towels, etc). The value of the gift is about 4.5 USD. If the most knowledgeable respondent is not present at the time of the visit, then the surveyor schedules a follow-up visit.

The second round of data collection occurs one year after the first round in each county. We collect data from the same households as in the first round, and were also able to extend the original sample by 10 randomly sampled households within the inner zone of the planned terminal location.[15] If either the survey respondent or the primary earner of the initially surveyed household no longer resides at the same address, we record this in our data and replace the household with another randomly sampled household within the same sampling zone (inner or outer). In our welfare analysis, we report results both before and after weighting each sampled household in proportion to the share of the village population in the sampling zone.

We collect detailed information about household consumption expenditures across 9 household consumption categories for retail products (food and beverages, tobacco and alcohol, medicine and health, clothing and accessories, other every-day products, fuel and gas, furniture and appliances, electronics, transport equipment) as well as for expenditures on business inputs. We also collect information about household incomes, hours worked and economic activities of different members (occupation (e.g., farmer, manual worker, etc.) and sector (agricultural, manufacturing, services)), in addition to data on asset ownership, financial accounts, internet use and migration.

Finally, in one of the 8 counties, the local government suspended further activity by our teams

[15] This extended sample was possible due to a small remaining positive balance on the project account that we decided to invest in expanding the household survey sample.

after we had completed endline data collection for 8 out of 12 villages in that county. This was unrelated to our operation, which followed the same protocol as elsewhere. As a result, we have endline data for 96 instead of the 100 villages. As the timing of data collection within the county was random, the 4 missing villages are not particular in any way. They include two control villages and two treatment villages.

Tables 1 and 2 and appendix Tables A.1-A.4 present descriptive statistics of the baseline data from the household survey at the individual level and at the household level. The tables also present descriptive statistics for the same outcomes in the control group at the endline data collection. The median age of all household members in the baseline survey is 44 and the median household size is 3. 60 percent of households report that the primary earner is a peasant, and 82 percent of households report that the primary earner completed at least primary school. In terms of demographics, these statistics are very similar to nationally representative rural household samples from the China Family Panel Study as well as the most recent Chinese Agricultural Census for the year 2016. The same holds for household economic characteristics: mean monthly income per capita and retail expenditure per capita are about RMB876 and RMB732 respectively, which makes these households significantly poorer than households living in urban centers. At baseline, households spend on average half of their retail expenditure outside the village, which requires travel as their main shopping destination outside the village is generally an urban center at a median two-way distance of 40 minutes. In terms of work location, 80 percent of primary earners work inside the village.As discussed in the introduction, many households report using the internet via smartphones or other devices: close to 40 percent report having used the internet, more than 50 percent own smartphones and close to 30 percent report owning a laptop or PC. Almost all households own a TV. At the same time, e-commerce penetration is very limited compared to urban regions: the average share of household retail expenditure on local e-commerce deliveries is less than 1 percent, and this does not change over time for the endline survey in the control group of villages. Similarly, the share of revenues from online selling in monthly household income is less than 0.5 percent, and again this does not change over time for the endline data collection in the control group of villages. By comparison, a recent survey conducted by McKinsey (2016) has found that urban households in Chinese cities spend on average up to 20-30 percent of total retail consumption on e-commerce deliveries.

Local Retail Price Survey Data

We aim to collect data on 115 price quotes for each village. 100 of these prices are from the same 9 household consumption categories for retail products as in our household survey (food and beverages, tobacco and alcohol, medicine and health, clothing and accessories, other everyday products, fuel and gas, furniture and appliances, electronics, transport equipment), and 15 price quotes are for local production/business inputs. Our protocol for the price data collection closely follows the IMF/ILO standards for store price surveys that central banks collect to compute the CPI statistics. The sampling of products across consumption categories is based on budget shares of rural households in Anhui and Henan that we observe in the microdata of the China Family Panel Study (CFPS) for 2012. The sampling across stores is aimed to provide a representative sample of local retail outlets (stores and market stalls). In villages with few stores we sampled all of them. The sampling of products within stores is aimed at capturing a representative selection of locally purchased items within that outlet and product group. Each price quote is at the barcode-equivalent level where possible (recording brand, product name, packaging type, size, flavor if applicable).

In the second round of data collection (one year after the first round), we aim to collect the price quotes of the identical products in the identical retail outlet where this is possible (see Appendix C). Where this is not possible, due to either store closure or absence of product in the store, we record the reason for the absence, and include a new price quote within the same product category that is sampled in the same way as in the first round.

Tables 2 and A.4 presents descriptive statistics of the baseline data from the local retail price surveys. Unsurprisingly durable goods categories (furniture and appliances, electronics and transport equipment) are an order of magnitude more expensive than goods in non-durable categories. The median number of sampled stores is 3 per village (40 percent of villages have 3 or fewer stores in total). These stores are small with a median floor space of $50m^2$, and the median store has not added any new product within the last month.

Firm's Admin Database

We complement the collected survey data with administrative records from two of the firm's internal databases that we access through a remote server. To the best of our knowledge, this is the first time that the firm has agreed to grant access to their internal database to external researchers. The first database covers 5 provinces (our three study provinces plus two additional provinces with high shares of rural population: Guangxi and Yunnan) over the period from November 2015 (1 month prior to the start of our survey data collection) to April 2017. This database covers the universe of e-commerce purchases made through the program in every participating village over this period. As summarized in Table 3, the purchase database covers approximately 27.3 million transaction records across 12,000 village terminals over the 18-month period. For each transaction, the database contains information about the terminal location, product category, number of units, amount paid and a unique buyer identifier.[16] Given that many terminals had already been in operation for several months prior to November 2015, these data cover adjustment periods beyond the 12-months window that we are able to capture as part of the RCT: terminals are observed up to two years and 4 months after the installation in these data. The second database covers the universe of sales transactions, i.e. out-shipments from the villages, through the firms distribution network for the same universe of roughly 12,000 village terminals in the 5 provinces over the period January 2016 to April 2017. For each transaction, the database contains information about the village of origin and the weight of the out-shipment in kg. As depicted in Table 3, the total number of e-commerce out-shipments over this period is roughly 500,000 . The table provides descriptive statistics for both datasets that we use in the analysis reported in Section 5.

Township-Level Data on Trade Market Access

We use geo-coded township-level data from the Chinese population census in 2010 in order to estimate the fraction of a rural location's total trade market access that stems from trading relationships with other rural locations in the same county, as opposed to access to larger urban

[16]We are able to identify 40 of the 43 e-commerce terminals in our RCT sample villages based on the Chinese county and village names that we have access to in the firm's transaction database.

markets within and outside the county. To do this, we use information on the recorded population[17] for each of roughly 45,000 township-level administrative units in China,[18] the coordinates of the centroid of each of those units, the type of township-level unit (e.g. urban districts, rural townships) and data on the value added per rural and urban worker at the province level for 2010. Sections 3.3 and 4.4 below provide further discussion and details about the estimation.

3 Theoretical Framework

This section proceeds in three parts. We first describe the channels through which the program can affect the local economy. We then derive a general expression of the program's effect on household economic welfare that guides the survey data collection and the empirical analysis in the following sections. Finally, we rationalize our empirical counterfactual in the light of potential GE spillovers across villages in the countryside.

3.1 Channels at Work

What type of economic shock does the program imply for the local economies? The program makes no investment in internet accessibility for villagers, and the terminal cannot be used to browse the internet except for the e-commerce platform. This, together with the fact that roughly 40 percent of village households report using the internet before the arrival of the program, and more than 50 percent own smartphones (Table 2), indicate that the shock is specific to the arrival of e-commerce, rather than providing internet access more broadly. Being able to separate the effects of e-commerce from first-time internet access more broadly (e.g. through emails, weather forecasts, online search, social media or online news) is one of the strengths of this empirical setting.

The program has two central elements that are aimed at removing the logistical and transactional barriers to rural e-commerce. First, the program aims to bring e-commerce-related shipping costs to and from the village to the same level as those present in the county's main urban centers. To this end, the program builds warehouses as logistical hubs for village deliveries and pickups, and fully subsidizes transport costs between the county's city center and the villages. Second, the program installs an e-commerce terminal in a central village location, where a terminal manager assists villagers to buy and sell products through the firm's e-commerce platform using traditional offline payments.

Both of these interventions affect the degree of trade integration between the village and the rest of urban China that is already connected to e-commerce. The logistical element reduces the physical trade costs to and from the village for bilateral pairs that are connected to e-commerce. At the same time, the program does not directly affect the transport costs of non-participating villages, or trade flows of program villages outside of e-commerce.[19] The transactional element (terminal installation) potentially reduces information and transactional frictions for trade flows to and from the village: e-commerce enables villagers to observe products and prices from other

[17] This includes both the registered and non-registered population currently residing in the unit at the time of the census.

[18] Townships are the most disaggregated unit of observation that we can obtain the full census database for. In China's administrative hierarchy, townships are one layer above villages. In the countryside, townships include on average about 14 villages. In urban regions, township-level units are small urban districts.

[19] As discussed in Section 2, the warehouses or distribution used for e-commerce transactions cannot be used for offline transactions outside the firm's platform.

regions that are connected to e-commerce far beyond the local economy and, in turn, other regions can learn about the products and prices from local producers. To the extent that villagers were already aware of the online information offered by the e-commerce platform in absence of the program (e.g. through smartphones), the terminal installation may still alleviate transactional barriers by making it easier for villagers to buy from or sell to trade partners outside the village.

By overcoming both logistical and transactional barriers to e-commerce integration, the program provides villages with essentially urban market access through e-commerce. In the majority of villages that were not previously served by commercial parcel delivery services, the effect that we observe will be driven by the removal of both of the barriers to e-commerce integration that we discuss above. In the fraction of villages that already were serviced by commercial parcel delivery distribution networks, there is in principle no pre-existing logistical barrier to e-commerce, and the comparison between treatment and control villages will be driven only by the additional provision of the terminal interface (removal of the second barrier to rural e-commerce).[20]

3.2 Quantifying Changes in Household Economic Welfare

As discussed above, the intervention that we are interested in evaluating has the potential to not just affect individual behavior and the nominal earnings of households, but also household cost of living in the denominator of real incomes. To empirically quantify the change in household price indices due to the arrival of the e-commerce program in the village, we require theoretical structure on the demand side.

Following existing work by Hausman (1996), Hausman & Leonard (2002) and more recently Atkin et al. (in press), we start with the compensating variation (CV) for household h. The CV captures the change in exogenous income required to maintain the initial level of utility in period 0 after the e-commerce program has arrived in period 1:

$$CV_h = \underbrace{\left[e(\mathbf{P}^1, u_h^0) - e(\mathbf{P}^0, u_h^0)\right]}_{\text{Cost of living effect } (CLE)} - \underbrace{\left[y_h^1 - y_h^0\right]}_{\text{Nominal income effect } (IE)}, \tag{1}$$

where $\mathbf{P}^t$ is the vector of prices faced by the household in period t, u_h^t is the household's utility and y_h^t is its nominal income.

The first term is the cost of living effect, the welfare change due to the price changes induced by the arrival of e-commerce. The second term is the nominal income effect, the welfare change due to any changes in household income that result from the arrival of e-commerce. While, at least in principle, we can record the effect on nominal household earnings and labor supply directly as part of the survey data collection, this is not the case for the cost of living effect. The store price survey data described above allow us to observe the vector of price changes $\mathbf{P}_C^1 - \mathbf{P}_C^0$ for continuing products in continuing local retailers (i.e. stores, market stalls, etc) that are present both before and after the arrival of the program. We index such continuing product prices by C.

However, there are three sets of price changes that are inherently unobservable: the consumer price changes $\mathbf{P}_T^1 - \mathbf{P}_T^0$ due to the entering e-commerce terminal indexed by T, the price changes

[20]The transport cost subsidy does not affect villages that were previously serviced by commercial parcel delivery services. The logistics operators offered service in a handful of rural locations at the same rate as elsewhere in the county prior to program entry. In those villages, households could order or sell online subject to pickup or delivery at an agreed central village location.

$\mathbf{P}_X^1 - \mathbf{P}_X^0$ of potentially exiting local retailers or varieties within continuing stores indexed by X, and the price changes $\mathbf{P}_E^1 - \mathbf{P}_E^0$ due to local store entry or new product additions in pre-existing local retailers. For example, prices at the new e-commerce terminal option cannot be observed in period 0, and exiting local retailers' prices cannot be observed in period 1. As first noted by Hicks (1940), we can replace these three unobserved price vectors with 'virtual' price vectors, the price vectors that would set demand for these shopping options equal to zero given the vector of consumer prices for other goods and services.

In the following, we denote such price vectors with an asterix (the implicit prices that would set consumption equal to zero in a given period), and break up the total consumption price vector in expression (1), into the four different components of potential consumer price changes. This leads to a decomposition of the program's total cost of living effect (CLE) into different channels that we can map to observable moments in the survey microdata:

$$\begin{aligned} CLE = & \underbrace{e(\mathbf{P}_T^1, \mathbf{P}_C^1, \mathbf{P}_X^{1*}, P_E^1, u_h^0) - e(\mathbf{P}_T^{1*}, \mathbf{P}_C^1, \mathbf{P}_X^{1*}, P_E^1, u_h^0)}_{\text{(1) Direct price-index effect } (DE)} + \underbrace{e(\mathbf{P}_T^{1*}, \mathbf{P}_C^1, \mathbf{P}_X^{1*}, P_E^{1*}, u_h^0) - e(\mathbf{P}_T^{0*}, \mathbf{P}_C^0, \mathbf{P}_X^{0*}, P_E^{0*}, u_h^0)}_{\text{(2) Pro-competitive price effect } (PP)} \\ & + \underbrace{e(\mathbf{P}_T^{1*}, \mathbf{P}_C^1, \mathbf{P}_X^{1*}, P_E^1, u_h^0) - e(\mathbf{P}_T^{1*}, \mathbf{P}_C^1, \mathbf{P}_X^{1*}, P_E^{1*}, u_h^0)}_{\text{(3) Entry effect } (EE)} + \underbrace{e(\mathbf{P}_T^{0*}, \mathbf{P}_C^0, \mathbf{P}_X^{0*}, P_E^{0*}, u_h^0) - e(\mathbf{P}_T^{0*}, \mathbf{P}_C^0, \mathbf{P}_X^0, P_E^{0*}, u_h^0)}_{\text{(4) Exit effect } (XE)} \end{aligned} \tag{2}$$

The first term of the first bracket and the second term of the final bracket of this decomposition represent the same difference in expenditure functions as in the first term of (1): the amount of expenditure one would have to pay household h in order to obtain the pre-terminal level of wellbeing, but evaluated at the post-intervention consumption prices. The terms in the middle between these two terms cancel out one another, so that this decomposition yields the total gains or losses due to effective consumption price changes, including changes at the extensive margin of consumer choice (e.g. new shopping options).

The first bracket, the direct price index effect, captures the consumer gains due to the arrival of the new terminal shopping option holding all other prices fixed. These gains can arise from three distinct channels that are all captured in the quantification of the bracket: e-commerce can provide existing products at cheaper prices, it can offer new product variety that was not previously available, and it can offer different shopping amenities (e.g. convenience, saving trips outside the village, etc). The second bracket captures changes in household cost of living due to price changes among pre-existing retailers and their products. For instance, existing retailers could lower their markups due to increased competition from e-commerce. Following Atkin et al. (in press), we label this the pro-competitive price effect. The third and fourth terms, that we refer to as the entry and exit effects, capture changes in product availability in the local retail environment. For example, the arrival of the e-commerce terminal could lead some local retailers to exit, it could in principle lead to store entry (e.g. stores sourcing online), and it could lead to disappearing or new product variety among pre-existing stores (for example due to local retailers starting to source their products online).

Up to this point, the welfare expressions in (1) and (2) are fully general, and do not rely on specific functional forms. However, the terms (1), (3) and (4) in equation 2 involve price index effects due to extensive-margin changes in consumer choice across retailers and products. Quantifying the implications of these unobserved price changes for household welfare requires imposing theoretical structure on the expenditure function. That same specification of consumer preferences will also provide a specific price index formula for the pro-competitive cost of living effect (2), which depends on observable prices.

The logic behind this approach is as follows: once we know the shape of the demand curve that governs consumer substitution across different retailer options within a given product group, we can use the observed changes in the expenditure shares across different shopping options before and after the program intervention in order to infer the unobserved effective consumer price changes that underlie this observed substitution. Once we know the elasticity subject to which consumers usually substitute across retailers as a function of differences in value-for-money, then one can back out the implied effective price index change for consumption of a given product group that is consistent with the observed substitution into the e-commerce terminal for that product group. Again, these price index changes could be driven by price differences, different product availability and/or different shopping amenities. The moments that inform the welfare evaluation are the observed changes in consumption expenditure shares in combination with knowledge about the consumer demand curve across retail outlets. A very similar logic follows when evaluating the price index movements due to disappearing stores, entering stores or product additions and disappearances within continuing stores.

In Appendix B, we outline one such approach to guide the empirical estimation that follows a nested CES preference specification commonly used in international trade and macroeconomics. In particular, local households are assumed to have Cobb Douglas perferences across broad product groups in retail consumption (durables and non-durables). Within these nests, groups of different household types have CES preferences across retailers (e.g. e-commerce terminal, stores or stalls in village, stores outside village, etc). Within stores, households choose across varieties on offer within product groups as a function of quality-adjusted product prices. This structure closely follows recent work by Atkin et al. (in press) on Mexican households, as we describe further in Section 6 below and in the appendix.

Regardless of the particular demand specification one imposes, the raw empirical moments that are required to quantify the welfare impact of the intervention fall into three different types. The first set of empirical moments are estimates of the causal effects of the intervention on a number of observable economic outcomes, such as the effects on household nominal incomes to capture the second term in (1), the fraction of total retail expenditure substituted into the new e-commerce terminal across different product groups and by household types, the effect on the price changes from continuing products in pre-existing retailers, and the effect on the propensity of store exit and entry, and of product entry and exit in the local retail environment.

The second type of required estimates are empirical moments from the baseline data collection, such as consumption shares across product groups. The third type of moments are estimates of demand parameters that govern the degree of consumer substitution across retailers and products. This latter set of parameters differ across different functional form assumptions on the demand side. Our Chinese empirical context, however, lacks the rich panel of consumer scanner data required to estimate these demand parameters. One advantage of the approach we outline in Appendix B is that it allows us to use recent estimates for households of very similar income ranges reported in Atkin et al. (in press), which to the best of our knowledge are the closest empirical estimates on the nature of retail demand and consumer substitution in an emerging market environment, such as China. In addition to tying our hands to existing estimates from the literature, we also report quantification results across a range of alternative demand parameters to document the sensitivity of the welfare estimates across a range of assumptions.

3.3 GE Spillovers

Our estimation exploits differences in outcomes between program villages and comparable control villages. This raises the question to what extent these differences may reflect spillover effects from treated villages on nearby control villages. The presence and strength of spillovers on e.g. local incomes or product prices is a priori unclear, and will depend on the degree of trade integration between villages in rural regions. If Chinese villages are small open economies whose market access is mainly determined by trade with urban areas, rather than by trade with other small rural markets, then the extent of spillovers could be muted. On the other hand, if trading with other villages in the countryside is an important component of villages' trade market access, then GE effects across villages could play an important role. In addition to spillovers driven by trade linkages between villages, it could also be the case that households in control villages use program terminals in nearby villages to access e-commerce.[21]

The extent of such spillover effects is an interesting empirical question for three main reasons. First, we are interested in estimating the effect of the program on the level of household welfare among villages that receive the e-commerce program. If the control group is indirectly affected, then an empirical specification exploiting the difference in outcomes between treatment and control villages no longer directly speaks to the program's impact on treated villages. Second, even after correctly adjusting for indirect effects on the control group, the presence of spillovers would also have implications for the external validity of the conclusions. In the current setting, only a fraction of the Chinese countryside in any given county is part of the program. If we wanted to inform policy-making on the welfare consequences of scaling up e-commerce access in rural China to a larger fraction of the countryside, then the presence of spillovers would imply that treatment effects depend on the scale of the program's roll-out. Third, the presence of spillovers would change our understanding of the aggregate implications of the program, either in its current form or when evaluating a scaled-up version of the program. That is, rather than focusing on the welfare effects on treated communities, we are also interested in the overall impact of the program among rural households as a whole. Here, knowing the extent of spillover effects allows us to compute the average effect of the program on rural households as a function of direct and indirect exposure to the program whose averages we can measure in the data (or simulate when scaling up).

In our empirical analysis, we begin by comparing economic outcomes in treatment and control villages, under the baseline assumption that rural-to-rural GE effects are negligible. We then proceed in two directions. First, we use a methodology close to Miguel & Kremer (2004) to investigate to what extent plausibly exogenous variation in exposure to nearby treatment villages

[21] Another possible source of spillovers in this setting are rural-to-rural migration flows for which we can test directly.

affects local economic outcomes conditional on the local treatment status of the village in question. Second, we use trade theory as a guidance and construct village-level measures of market access. Market access is the weighted sum of access to market expenditures across all rural and urban market places in China and beyond, where the weights are inversely related to the bilateral trade costs on each potential trading route. We can use information on the geographical position and market size of all rural and urban settlements in China prior to the program's roll-out in combination with measures of bilateral travel costs in order to investigate what fraction of trade market access in our village sample is driven by access to urban markets relative to other villages within the same county that participate in the e-commerce expansion program. We implement these two approaches in Section 4.4 below.

4 Empirical Analysis Using Survey Data

In this section, we estimate the program's effect on a number of economic outcomes related to household consumption, incomes, economic activity and local retail prices, that we observe in the survey microdata. In addition to being of interest in their own right, these empirical moments enter the quantification of changes in household economic welfare in Section 6.

4.1 Average Program Effects

Following e.g. McKenzie (2012), we run the following regressions:

$$y_{hv}^{Post} = \alpha + \beta_1 Treat_v + \gamma y_{hv}^{Pre} + \epsilon_{hv}, \tag{3}$$

where y_{hv} is an outcome of interest for household h living in village v. For outcomes from the retail price data, h indexes individual price quotes or store-level outcomes instead. $Treat_v$ is an indicator of intended treatment according to our randomization, so that β_1 captures the intent-to-treat effect (ITT). We also estimate the treatment-on-the-treated (TOT) after instrumenting for the actual treatment status using $Treat_v$. Finally, we run (3) after replacing the binary treatment indicator with a continuous measure of the log of household residential distance to the nearest program terminal, again using $Treat_v$ as an IV.

For households who were either replaced or added as part of our extended sample in the second round (from 28 to 38 households), we define y_{hv}^{Pre} as the mean pre-treatment outcome of households living in the same zone (inner or outer) in the same village. The implicit assumption is that households were not induced to move within or across villages as a result of the program.[22] We cluster standard errors at the level of the treatment (village-level).

Tables 4-6 present the estimation results for the average effects on household consumption, incomes and local retail prices. Our discussion focuses on the TOT results, while the tables display the three types of effects discussed above (ITT, TOT and log distance). We run these regressions on the survey sample of households, stores, and price quotes described in Section 2 and Appendix C. For the welfare quantification in Section 6, we will also report results after re-weighting village zones according to their village population shares.

[22] As reported in appendix Table A.5, we find no evidence that households in treated villages are more or less likely to reside at the same address in the post-treatment survey. We also find no treatment effect on migration decisions of members within households.

Consumption

In Table 4, we find that the program on average leads to an uptake of 9 percent of households in treatment villages who report to have ever used the terminal for making purchases, relative to households in control villages. This treatment effect is about 5 percent when restricting attention to terminal use over the month prior to our endline survey.[23] These effects on consumption-side uptake may in part mask additional uptake from households in nearby villages. We return to this issue when investigating spillover effects in Section 4.4 below. The effect on the e-commerce terminal share in total household retail expenditure is 1.24 percent for the average household in our survey data. Thus, households that report ever having used the terminal spent on average 0.0124/0.089=14.1 percent of their retail consumption at the terminal during the past month. For those who bought over the past month, this share rises to 0.0124/0.049=25.3 percent.

Looking at retail consumption across product groups, we find stronger effects on durables compared to non-durables. For durables, the treatment effect on the terminal share of household expenditure is 6.7 percent for the average household in our sample, indicating a 44 percent shift in durable consumption to the e-commerce terminal among households who report to ever having used the terminal for consumption.[24] For non-durables, the treatment effect on the terminal share in household retail expenditure is 1 percent for the average household, indicating that ever-users spend on average about 11 percent of total non-durables expenditure at the terminal.[25] In contrast, we find no significant substitution to the e-commerce terminal for household expenditures on production and business inputs (e.g. fertilizer, tools, machinery, materials, etc). Finally, while households do shift part of their expenditures to the terminal, there are no significant treatment effects on total monthly retail expenditures. This result is consistent with the lack of income effects of the program that we discuss in the next subsection.

To summarize, the program leads a minority of local households to take up the new e-commerce shopping option. Among users, we find sizable effects on the substitution of total household retail expenditure to the e-commerce terminal, especially for durable consumption. These results are indicative of significant direct consumption gains for certain groups of local households. We return to the welfare computations based on these moments in the final section below.

Incomes

The income effect of e-commerce on local producers could be in principle either positive due to the possibility of selling online, or negative due to increased competition from the new terminal shopping option. In Table 5, we find no treatment effect on household incomes, or on labor supply as measured by hours worked by the primary and secondary earner. The point estimates on incomes per capita are close to zero and negative, and not statistically significant. We find no effects on either annual or monthly income, from agricultural or non-agricultural sources. In contrast to

[23]Following standard protocol, we construct monthly consumption based on the last two weeks of expenditures for non-durables (multiplied by two), and on the past three months for durables (divided by three). Usage over the past month is thus defined as either having purchased non-durables over the past two weeks, or as having purchased durables over the past three months. Appendix C.4 provides additional details.

[24]To compute durable consumption shares, the sample is restricted to households who buy any durables over the past three months. In this sample, the treatment effect on ever using the terminal is 15.3 percent instead of 9 percent. This yields an effect on the average durables consumption share among uptakers of 0.067/0.153=44 percent.

[25]Since all households consume non-durables, the treatment effect on uptake is as reported in Table 4, so that the average non-durables terminal share among ever users is 0.01/0.089=11 percent.

our consumption results, we find no treatment effect on online selling activity, online revenues or business creation offline or online. The point estimate on whether "any member of the household has ever sold online" is also close to zero, negative and not statistically significant. Given that the control group experienced no increase in income shares from online selling activity relative to its tiny level at baseline over this period (Table 2), these estimates suggest that the e-commerce program had no significant effect on the uptake of online selling activity or local revenues.

We are cautious in drawing conclusions on the absence of production treatment effects from our household survey only. The 12-month period between baseline and endline surveys may be too short for local households to grow their online selling activities. Our survey sample may also fail to capture rare but highly successful tail events of online businesses within villages that could shift the local mean effect on incomes. Furthermore point estimates on nominal incomes that are based on survey data are noisy. In Section 5 we use the firm's internal database to corroborate our analysis that is based on the survey data. As discussed below, these admin data allow us to observe the universe of buying and selling transactions, and to estimate the monthly time path of adjustments both before and after 12 months post-program entry.

Local Retail Prices

Table 6 shows the average program effects relative to control villages using the retail price survey data. We find no significant reduction in local store prices for identical continuing products that we observe in the same local retailer in both baseline and endline data. The point estimate is close to zero and positive, and not statistically significant. Given our sampling framework in Section 2, the unweighted average effect on local retail prices resembles the effect measured by a Laspeyres price index for local retail consumption.[26]

One piece of evidence suggests potential knock-on effects on pre-existing local stores. The treatment effect on the number of new products per store over the past month is 4 goods and significant at the 10% level. This positive effect is large relative to the mean number of new goods of 1.4 in the baseline (2), but still small relative to the expected stock of goods in stores.[27] Furthermore, we find a negative but statistically insignificant effect for durable products. Given the small sample size of durables observed in the villages, this could be consistent with the more pronounced treatment effect on household durable consumption above. We re-visit the plausibility and robustness of these knock-on effects on local stores in the heterogeneity analysis that follows below. Finally, we should again point out a limitation to the scope of our survey data collection: while we are able to estimate pro-competitive effects on the local retail price environment within the villages, potential effects on retail prices in nearby urban centers, where households source part of their consumption, would be outside the scope of these data. Given how small villages are compared to urban centers within counties (see Section 4.4 below), and the fact that only a small fraction of all villages participate in the program during our sample period, GE effects on urban centers are somewhat unlikely in our current setting. Having said this, and following the discussion in Section 3.3 above, potential GE effects on prices and incomes in urban China could

[26]Our price survey data collection follows the data collection protocol of the IMF data dissemination standard for CPI analysis across countries. For example, the BLS in the US or INEGI in Mexico estimate a Laspeyres price index across product groups using a similar methodology.

[27]We find no significant effect on store online sourcing, but this average appears to mask significant heterogeneity with respect to the initial availability of commercial parcel delivery. We return to this result in the next section.

play a role in the future, when evaluating the scaling up of e-commerce expansions to larger parts of the countryside.

4.2 Heterogeneity Across Households and Villages

We now investigate the extent to which the average effects mask significant heterogeneity across households and villages. To this end, we estimate regressions of the following form:

$$y_{hv}^{Post} = \alpha + \beta_1 Treat_v + \beta_2 X_{hv} + \beta_3 Treat_v \times X_{hv} + \gamma y_{hv}^{Pre} + \epsilon_{hv}, \tag{4}$$

where X_{hv} indicates different pre-existing household or village characteristics. As before, we report the results of specification (4) for both ITT and TOT, and after replacing the binary treatment variable with log household residential distance to the nearest terminal location (again instrumenting with planned treatment status). We begin by investigating the heterogeneous effect of the program with respect to pre-existing availability of commercial parcel delivery at the village level. Villages that were already serviced by commercial parcel delivery operators during our baseline survey were essentially already connected to the same e-commerce logistical network as urban centers in the same county prior to the program's arrival.[28] Interacting the treatment with pre-existing parcel delivery status therefore allows us to shed light on the effect of removing both the transaction and logistical barriers to rural e-commerce (among villages without pre-existing parcel delivery), from the effect of removing only the transactional barrier (in villages with pre-existing parcel delivery). Next, we investigate heterogeneity across a basic set of pre-existing household and village characteristics: respondent age, education, household income per capita, residential distance to the planned terminal location, and a measure of village remoteness based on road travel distance to the nearest township center.

Table 7 reports the heterogeneous impact of the program with respect to pre-existing commercial parcel delivery across a number of economic outcomes. On the consumption side, we find that the average treatment effects are driven by villages that were not initially connected to commercial parcel delivery services. The average effects among the previously connected villages are relatively precise zeroes on all outcomes that showed significant average effects in the pooled sample. This somewhat magnifies the previously reported average treatment effects on terminal consumption in the 85 percent of the village sample not previously connected to commercial parcel delivery. In these villages, slightly more than 10.5 percent of local households are induced to ever use the terminal relative to the control villages, and the average household spends 1.5% of their total retail expenditure on the e-commerce terminal over the past month. On the production side of the local economy, however, we find no significant effects in either group of villages, confirming the earlier pooled results. Considering the local retail outcomes, we now find a significant treatment effect on the number of stores sourcing their products online in villages without pre-existing delivery, and again find a treatment effect on new product varieties that is significant only in these villages. The treatment effect on local durable prices increases to -14.4 percent in villages without pre-existing delivery, but remains statistically insignificant at conventional levels. These results suggest that the removal of the logistical barrier to e-commerce is the main driver of

[28] As discussed in Sections 2 and 3, the logistics operators offered service in a handful of rural locations at the same rate as elsewhere in the county prior to program entry. In those villages, households could order or sell online subject to pickup or delivery at an agreed central village location.

the program's local economic effects, rather than the provision of an additional terminal interface in villages that already have a logistical connection to e-commerce.

Table 8 extends the analysis of heterogeneous treatment effects to other household and village characteristics. We first run regressions in which only one characteristic is interacted with the treatment, and then we run a combined regression with all interactions included jointly. We find that younger, richer households who live in closer proximity to the planned terminal, and in villages at longer distances from the nearest city center experience significantly more positive treatment effects on the consumption side. In particular, Table 8 shows that the average effect on terminal uptake is driven by, and more sizable among these groups of households. Somewhat surprisingly, we find no significant heterogeneity in household usage of the terminal with respect to the education (years of schooling) of the household respondent. And we again find no significant heterogeneity in the treatment effect on the production side of the local economy, or on pro-competitive price effects among local retailers. We return to the heterogeneity of the program's effects as part of the welfare analysis in Section 6.

4.3 How Does E-commerce Compare to Pre-Existing Shopping Options?

Before providing additional evidence from the firm's admin database in the next section, we use the survey data to investigate the roles of features of the program implementation and spillovers on control villages underlying the observed effects. In this subsection, we use the survey data to describe how the arrival of e-commerce compares to the pre-existing retail environment of local households. Table 9 reports a number of descriptive statistics. Overall, the new e-commerce terminal compares favorably with pre-existing shopping options on a number of dimensions, including accessibility, value-for-money and product variety.

To illustrate the pre-existing retail environment, recall from Table 2 that households source more than half of their total retail consumption outside their village, and almost 70% of their durable goods consumption. The need to travel outside the village to shop is unsurprising, given that our surveyors could not find any durable goods in local stores for about half of our sample villages (Table 9). Households main reported shopping destination outside the village is at a median distance of 10 km return trip, representing a 40 minute round trip at a median cost of 4 RMB (Table 9).

In comparison, the terminal is much closer to our survey households, with a median distance to the planned terminal of 230 m (Table 2). The terminal also offers a variety of goods unavailable in local stores. As shown in Table 9, 62 percent of goods bought through the e-commerce program were not available in the village, which rises to 84 percent for durable goods. When goods are available at both the terminal and in the village, the terminal is cheaper by a median price reduction of 15 percent.[29] The main shopping destination outside the village, generally the nearest township center, is more competitive in terms of varieties offered (80 percent of goods purchased on the terminal are available there), but the terminal remains cheaper by a median of 18 percent even before accounting for transport costs. Given fast processing at the warehouse locations, e-commerce delivery times in program villages are close to identical to those in urban regions within the county.

[29] As part of our survey, we elicit for each e-commerce purchase, whether the good was available in the village. If the good is available, we ask how much it would have cost.

Despite these advantages, not all rural households substitute expenditure into the new shopping option. In this context, poor program implementation could be an explanation for why the program did not attract a broader cross-section of the local population. This seems a priori unlikely, given the firm's high degree of professionalism, profit motive, institutional capacity and expertise, especially when compared to the resources generally available to implement public policies in developing countries. To further investigate the importance of program implementation in explaining household take-up, we obtain information about features of program implementation across villages. In particular, we observe the terminal manager application test score, and a dummy for delay in the terminal installation with respect to the planned due date in the implementation schedule. We find that neither of these program features affect take up of the terminal in a significant way. These results and the general context of the intervention both point against the possibility of a botched program implementation that significantly affect the survey results.

4.4 Role of Spillovers

We next investigate the role of GE spillovers on surrounding villages that could in principle confound our findings from the survey data, as discussed above in Section 3.3. For example, if trade linkages with other nearby villages are an essential driver of the local economy, then it could be the case that the comparison between treated and control villages misses average income effects. If these villages are well integrated with one another, it could also be the case that store prices in surrounding villages respond to pro-competitive effects, potentially biasing toward zero the comparison between treatment and control villages. To investigate these mechanisms, we pursue two different approaches.

First, we follow an approach similar to Miguel & Kremer (2004):

$$y_{hv}^{Post} = \alpha + \beta_1 Treat_v + \beta_2 Exposure_v^{treat} + \beta_3 Exposure_v^{all} + \gamma y_{hv}^{Pre} + \epsilon_{hv}, \tag{5}$$

where $Exposure_{vk}^{treat}$ measures the proximity of village v to other program villages, and $Exposure_{vk}^{all}$ measures proximity to all villages on the candidate list from which we randomly selected our control villages. Even though exposure to other program villages is not randomly assigned, our randomization means that conditional on exposure to all candidate villages, exposure to other treatment villages is plausibly exogenous. In turn, β_2 is an estimate of the the strength of cross-village spillovers. We measure exposure as the number of intent-to-treat villages within 3 or 10 km distance bins of a given village. Table 10 reports the estimation results. We find some evidence of positive spillover effects of nearby terminals within 3 km of the village. These effects imply a larger total average effect of the program installation on household uptake that we estimated above. This increases from 9 percent in Table 4 to about 14 percent once we take into account positive spillovers from nearby villages, and about 13 percent in the village population when adjusted for sampling weights. In contrast, we find no evidence of cross-village spillovers on local retail stores, or on the production side of the economy.

To further investigate these channels in the absence of experimental variation in program saturation rates,[30] we also pursue an approach grounded in trade theory. In particular, we can

[30] As part of our negotiations and collaboration with the firm's local implementation teams, it was not feasible to also attempt a two-stage cluster randomization design that would have allowed us to randomly vary saturation rates.

quantify the fraction of a rural location's total trade market access that is due to trading exposure to other rural markets in the same county. This fraction provides additional information on the extent of rural-to-rural spillovers from other sample villages in our setting. If a sizable share of local market access is due to trading relations with other local rural markets, then indirect effects on local product prices and incomes from treatments in other villages could become an important force. If, on the other hand, local product and factor prices are predominantly determined by access to larger urban markets, then rural-to-rural spillovers could have negligible effects on local prices and incomes across our sample villages.

Following e.g. Head & Mayer (2014), the market access of location v to all other rural and urban markets $j \neq v$ is:

$$MA_v = \sum_{j \neq v} \tau_{jv}^{-\theta} \, Y_j \tag{6}$$

where τ_{jv} is the bilateral trade cost, θ is the elasticity of trade flows with respect to trade costs, and Y_j is a measure of j's market size.[31] MA_v is thus a weighted sum of economic activity outside of market v, with weights that are inversely related to bilateral trade costs. To compute the fraction of total market access that is due to bilateral linkages with other rural markets in the same county (i.e. MA_v^R / MA_v), we compute (6) both across bilateral connections to all other markets (denominator), and only summing across bilateral connections with other rural markets in the same county (numerator). Alternatively, we restrict the numerator to bilateral connections with respect to the fraction of rural markets in the county that are participating in the program to compute the share of market access due to rural locations with program terminals. That fraction was about 1/6th of all rural markets in participating counties over our sample period.

To compute these measures, we use the township-level data from the Chinese census in 2010 described in Section 2. These data provide us with the populations residing in each of roughly 45,000 township-level administrative units. In addition, we use the coordinates of their centroids to construct the full matrix of bilateral distances in km. Following the trade literature, we use these bilateral distances to parameterize $\tau_{jv}^{-\theta}$: using the finding that the elasticity of trade flows with respect to distance is approximately -1,[32] we measure $\tau_{jv}^{-\theta}$ as the inverse bilateral distance in km when summing across the j market sizes. Alternatively, we also use a larger distance elasticity of -1.5 that gives more weight to markets in closer proximity. For market size Y_j, we use populations or populations multiplied by the value added per worker for rural and non-rural workers measured at the province level for 2010. The first metric provides an inverse distance-weighted measure of market access to populations outside the township, while the second provides an approximate measure of access to GDP. Finally, we define rural and urban markets following the administrative classification across township-level units we obtain in the census data. For computational feasibility, when constructing the full matrix of bilateral connections, we compute the total market access of rural townships with respect to all other township units (both rural and urban)

[31] To be consistent with structural gravity in trade models, the measure Y_j of j's market size would include a multilateral resistance term capturing j's own degree of access to all other markets (see e.g. Head & Mayer (2014)). In (6), we abstract from this and compute a first-order approximation of the structural gravity expression for MA_v. In practice, these measures have been found to yield very similar results in recent empirical work, as they are highly correlated (e.g. Donaldson & Hornbeck (2016)).

[32] See e.g. Disdier & Head (2008) for a meta-analysis of this point estimate.

within each of the 3 broad administrative regions of China in which our sample counties are located: East China (7 provinces), Middle China (3 provinces) and Southwest China (5 provinces).[33]

The above provides us with four measures of the ratio of total market access that is due to access to other rural populations or rural GDP within the the same county: measured either in terms of access to population or to GDP, and measured either in terms of access to all rural markets in the county or only the fraction of rural markets that on average participate in the e-commerce program. We compute the median, mean and standard deviations of these 4 ratios for all rural townships located in the three regions of China, as well as only for townships in our 3 sample provinces, or only for townships in the 8 sample counties. Furthermore, we compute each of these measures both for the baseline distance elasticity of -1, and when using -1.5 instead.

Appendix Table A.7 presents the estimation results. Overall, we find that other rural markets in the same county account for a tiny fraction of total trade market access for the median or the average rural market place. This result is driven by the fact that nearby rural markets within the same county account for a small fraction of the market size that is concentrated in vastly larger urban centers. This is particularly the case when using economic output as the measure of market size, but also holds for raw populations. For example, the median fraction of market access from nearby rural markets in terms of GDP is 0.37 percent in our sample provinces, and 1.2 percent in terms of population access. These fractions slightly increase when giving more weight to nearby markets using a higher distance elasticity, but remain close to zero in both cases when computing rural-to-rural market access only with respect to the average fraction of rural markets that are participating in the program in any given county over our sample period. These findings are in line with the absence of significant GE spillover effects on market prices or nominal incomes from our first approach above, and serve to provide some further corroborating evidence in this context.

Summary of Findings from the Survey Data

We can summarize the results of this section as follows. On the consumption side, we find that the program leads to sizable substitution of retail expenditure among households who are induced to use the new e-commerce terminal shopping option. These households represent about 14 percent of the rural household sample and about 13 percent of the village population after adjusting for sampling weights. We find that the program's effect is subject to significant heterogeneity. The beneficiaries are on average younger, richer, live in closer proximity to the program's terminal and in villages that are more remotely located. Conditional on these characteristics, we do not find evidence that household education or the characteristics of the terminal manager are significant determinants of the program's impact. The consumption response is mainly driven by the removal of the logistical barrier in villages with no pre-existing commercial parcel delivery, rather than by lifting additional transactional hurdles through the terminal interface. The new e-commerce option offers on average cheaper prices, more product variety and convenience/less travel costs. We find that the consumption effects are particularly pronounced for durable product groups, such as electronics and appliances. We also find suggestive evidence of pro-competitive effects on the local retail environment: local store owners report significantly higher numbers of new product variety, and a higher likelihood of sourcing their products on-

[33] Omitting provinces outside these zones is somewhat conservative, as their inclusion would increase the denominator of the rural-to-total market access ratios.

line in treated villages who did not initially have commercial parcel delivery. We do not find significant price reductions among local stores. On the production side, we find no evidence of significant effects on the local economy in terms of online selling activity, purchases of business inputs, household incomes, labor supply or entrepreneurship.

5 Additional Evidence Using the Firm's Admin Database

In this section, we use the firm's internal transaction database to provide additional evidence on four remaining questions that are outside the scope and budget of our household survey data collection. First, are the villages in our RCT sample representative of villages targeted by the program across the Chinese countryside more broadly? Second, to what extent does seasonality and the timing of our endline data collection affect the estimation results? Third, what is the time path of adjustments on the consumption and production sides, and is terminal take-up increasing beyond our survey's 12-month post-treatment time window? And fourth, is our survey data missing rare but highly successful tail events on the production side that could shift the average effect on local household income per capita?

As described in Section 2, we have access to the universe of purchase transaction records over the period November 2015-April 2017, across roughly 12,000 participating villages that existed over this period in 5 provinces. To capture household sales through the e-commerce terminals, we also obtained access to data on the universe of village out-shipments and their weight in kg for the same terminal locations over the period between January 2016 to April 2017.

Are the RCT Sample Villages Representative?

One concern is that the 8 counties that our RCT takes place in may not be representative of program villages in the Chinese countryside more broadly. To assess whether the RCT villages are representative of the population of program villages in China, we use the 5-province transaction database on both purchases and sales transactions to estimate regressions of the following form:

$$y_{vm} = \theta_m + \beta \text{RCTSample}_v + \gamma \text{MonthsSinceEntry}_{vm} + \epsilon_{vm},$$

where v indexes village terminals and θ_m is a set of monthly dummies indexed by m for the 18 months of operation from November 2015 to January 2017. y_{vm} is one of five terminal-level outcomes (monthly number of buyers, number of purchase transactions, total terminal sales, number of out-shipments and total weight of out-shipments in kg), *RCTSample* is a dummy for whether the terminal is in our RCT sample, and *MonthsSinceEntry* controls for the number of months that terminal v has been in operation as of month m. The standard errors ϵ_{vm} are clustered at the terminal level.[34]

The results in appendix Table A.8 show no remarkable differences between our RCT villages and the population of program villages in these 5 provinces. The same is true if we compare our RCT villages to all villages in our 3 survey provinces. The RCT sample seems marginally more successful on the out-shipment side, but the magnitudes are tiny. These results provide some reassurance against the potential concern that the e-commerce firm directed our team towards 8 counties that systematically differ from the program's target locations in the Chinese countryside.

[34] With very rare exceptions there is only one terminal per village.

Did We Collect Endline Data During Particular Months?

The timeline of pre-treatment data collection was determined by the roll-out schedule of the e-commerce firm, and we could not finance more than a single post-treatment round. As a result of these constraints, our survey cannot measure the impact of seasonality on treatment effects. We therefore use the transaction database to study seasonality effects by estimating:

$$y_{vm} = \theta_v + \beta \text{RCTMonth}_m + \gamma \text{MonthsSinceEntry}_{vm} + \epsilon_{vm},$$

where *RCTMonth* is a dummy for our survey months i.e., a dummy equal to 1 if month m is either in December, January, April or May, which are the four calendar months during which we conducted our survey. We again cluster standard errors ϵ_{vm} at the terminal level. The results are in appendix Table A.9. We find slightly higher numbers of terminal buyers during survey months relative to the rest of the calendar year, and slightly lower numbers of purchase transactions and out-shipments. In both cases, the point estimates are very small: about one additional buyer per month, 4-5 less monthly transactions, and a reduction of less than one out-shipment per month on the selling side. We conclude that seasonality is unlikely to be a significant driver underlying the findings of the RCT.

What Is the Time Path of Adjustments in Consumption and Production?

The program's objective to introduce e-commerce to all promising Chinese villages and continuous roll-out in our RCT counties imply that we cannot keep our control group untreated for more than one year. The firm's transaction data allows us to see beyond this one-year survey horizon, and to plot the time pattern of monthly terminal usage for both purchasing and selling starting from program installation. In particular, these plots tell us whether the impacts of the e-commerce terminals grow stronger over time, either on the consumption or production sides.

We estimate the following event study specification:

$$y_{vm} = \theta_v + \delta_m + \sum_{j=-3}^{24} \beta_j MonthsSinceEntry_{jvm} + \epsilon_{vm} \tag{7}$$

Each observation in equation 7 is a terminal in a given month. A negative index j denotes the number of months prior to installation for terminal v and in this case the outcome y_{vm} will always be 0. A positive value of j indexes the number of months since terminal v started operation, so that β_0 is a measure of average outcomes for terminals during the month of their installation, β_1 captures averages one month after installation, and so on. We assign an index of j =24 to all observations equal or beyond 24 months after the first month of program entry, so that β_{24} captures average outcomes of terminals that have been in operation for more than two years. Since we have terminal and month fixed effects, each of the β_0-β_{24} are estimated relative to the omitted category that are periods pre-installation (zeros by construction since the terminals did not exist).

To estimate (7), we create a balanced panel in the sense that each of the roughly 11,900 village terminals ever observed in the raw data appears once per month in the panel, for each of the 18 months for which we have data (16 months in the shipment data). This panel starts in November 2015 for the purchase database and in January 2016 for the out-shipment database. It spans terminal observations of up to 17 months pre-installation for villages connected in April 2017, to 28 months post-installation for the earliest terminals connected 10 month prior to the beginning of our data in November 2015.

In terms of identification, we no longer have experimental variation and a clear counterfactual control group when using the firm's internal database, as we did in the RCT. Instead, the assumption is that online purchases or out-shipments would be a hard zero in these villages if the program had not arrived in month $j = 0$. This assumption is reasonable given that online purchases or sales remain close to zero at endline in the control villages (Table 2). Reassuringly, we also find that the magnitudes of the program's effect after 12 months are closely aligned with the findings based on the RCT's survey data. On the other hand, if for some reason we believe this assumption not to hold in the broader set of villages that we are able to observe in the transaction data, then the estimates of the findings of the event study we discuss below can be interpreted as upper-bound estimates of the effect of the program (assuming a hard zero for the counterfactual).

Figures 2 and 3 present the event-study plots for terminal-level outcomes on the consumption and production sides. On the consumption side, we find little evidence of increasing uptake past our survey's one-year timeline. Terminal usage increases rapidly for about 2-4 months after opening, and then plateaus at around 80 buyers and 280 transactions per month per terminal. At the same time, total terminal sales in RMB appear to slightly decline over time, after peaking at about 3 months after program entry, suggesting that villagers make higher-value purchases first and then switch to buying lesser-value products through e-commerce.

On the production side, we find evidence that village-level number and total weight of out-shipments increase smoothly over time after program entry, and that this increase continues beyond the 12-month window that we cover in our survey data collection. The effect increases by roughly 50 percent when comparing the point estimate on the total weight of out-shipments 12 months post-entry to the point estimate for more than 2 years post-entry (including periods up to 2 years and 4 months post-entry). These results suggest that production-side adjustments may take longer to fully materialize than the 1-year horizon covered in the survey data. Despite this positive trend, the estimated effects at the village level remain relatively minor even two year post implementation. The average number of monthly out-shipments is about 10 in periods more than 2 years after the arrival of e-commerce. In turn, the combined weight of all village-level out-shipments increases to about 30 kg on average.

Are the Survey Data Missing Successful Tail Events on the Production Side?

Our survey sampling of 38 households per village may be insufficient to capture rare but very successful events on the production side. If neglected, such tail events of high-volume online businesses enabled by the terminal could in principle shift the average effect of the terminal on household incomes that we estimate as part of the RCT analysis. To investigate this issue, we use the universe of e-commerce shipments from 5 provinces over the period January 2016 to April 2017. As discussed above, we observe total shipment weight in kg, but not revenues. Figure 3 shows that the mean monthly number of e-commerce shipments out of the villages peaks around 10 with a mean total weight of less than 30 kg for the entire village.

To obtain a non-conservative upper-bound for these shipments' value to the local village economy, we assume that i) all of these shipments are pure local value-added and thus 1:1 adding to local incomes per capita, and ii) that the average value per kg of these shipments is as high as that of Chinese exports to the world (i.e. on average RMB66.5 per kg in 2015 and 2016).[35] Under these assumptions, we find that e-commerce out-shipments account for on average at most a 0.17 percent increase in local income per capita more than 2 years after the program's arrival. In summary, this upper bound of the average longer-term effect that we can estimate precisely in the administrative transaction data would still be consistent with the statistical zero result that we find using the survey data after one year in the RCT data collection.

Summary of Findings from Transaction Database

When comparing our RCT villages to the roughly 12 thousand other villages in the transaction data, we find that they are broadly representative of the Chinese village population that is being considered by the firm to be part of the e-commerce expansion program. The periods during which we collected endline data appear to be slightly above-average for some outcomes related to terminal purchasing use, and slightly below-average for some outcomes related to purchasing price tags and village out-shipments. However, the point estimates are very small in magnitude, suggesting that seasonality is unlikely to be a major factor in the RCT analysis. In terms of time path of adjustment, we find little evidence on the consumption side that the program's effect takes longer to materialize than the one-year period covered by our survey. The effects occur within 2-4 months after installation and remain roughly stable afterward. On the production side, we find evidence that village-level out-shipments are increasing significantly over time after installation. The effects remain small, however, in terms of total out-shipment weights, suggesting a minor upper-bound effect on village income per capita more than 2 years post-installation. Related to this, we find no evidence that our survey data collection missed rare but highly successful tail events on the production side that could have in principle shifted the village-level average effect on economic outcomes.

6 Quantification

This section combines the empirical results from the previous sections with the theoretical framework in Section 3 and Appendix B to quantify the program's effect on average household welfare, decompose the underlying channels, and estimate the distribution of the gains from e-commerce integration across households and villages.

Average Gains

The most robust evidence of significant treatment effects from the program from the previous sections is on the substitution of local households' retail expenditures to the new e-commerce terminal shopping option. As discussed in Section 3, these treatment effects enter the direct price index effect as part of the consumer gains from the program.[36] Even though it is impossible to di-

[35]From the World Bank's WITS database, which provides total value of Chinese exports and total weight.

[36]We also find suggestive evidence of additional product variety among local retailers. We abstract from this effect in the quantification for two main reasons. First, the point estimates are small as a fraction of the available product space (2 products after removing both barriers in Table 7). Second, quantifying the welfare implications of this effect would require a number of additional moments that are outside the scope of the fieldwork, such as knowing the local market shares of each of the barcode products and estimating an additional elasticity of substitution between products within a given retailer (the lower tier in Appendix B). Rather than imposing strong assumptions, we quantify a conservative estimate of the gains from e-commerce in this section, that could have been very slightly higher after

rectly observe the implicit price index changes due to the arrival of a new retail shopping option –that includes differences in prices, product variety as well as shopping amenities–, we can use existing estimates of the slope of household demand across retail shopping options to quantify the change in consumption value that is consistent with the changes in household expenditure that we observe in the data.

Following Feenstra (1994) and more recent work by Atkin et al. (in press) on Mexico, we derive an expression for the direct consumer gains from the arrival of the e-commerce terminal, expressed as a percentage of initial household expenditure:

$$\frac{DE}{e(\mathbf{P}_T^{0*}, \mathbf{P}_C^{0}, \mathbf{P}_E^{0*}, \mathbf{P}_X^{0}, u_h^0)} = \prod_{g \in G} \left(\Big(\sum_{s \in S_g^C} \phi_{gsh}^1 \Big)^{\frac{1}{\sigma_g - 1}} \right)^{\alpha_{gh}} - 1, \tag{8}$$

where $\sum_{s \in S_g^C} \phi_{gsh}^1$ is the share of retail expenditures that is not spent on the new e-commerce terminal post-intervention ($s \in S_g^C$ indexes continuing local retailers), σ_g is the elasticity of substitution across retail options to source consumption in product group g, and α_{gh} is the initial expenditure share on that product group for household group h. Appendix B provides a detailed derivation.

To estimate this expression empirically, we require information about the program's effect on $\sum_{s \in S_g^C} \phi_{gsh}^1$, as well as the parameters α_{gh} and σ_g. For the α_{gh}, we use our baseline data on household expenditure shares across product group. For ex-post expenditure shares on the new e-commerce option, we use the treatment effects among the 85 percent of villages without pre-existing parcel delivery connections reported in Table 7. These villages experienced the removal of both logistical and transactional barriers to e-commerce integration, which is the counterfactual that we focus on for the quantification exercise. We include the regression intercept (mean program usage among control villages) in these treatment effects on household terminal consumption shares to account for the positive spillovers that we estimate in Table 10.

We perform this welfare computation for two different groups of local households. First for the average sample household, for where the average treatment effect on the terminal share of total retail consumption is 1.6 percent, and second for households who report ever having used the terminal for consumption, for whom the average effect on the terminal expenditure share is 14 percent. Given the heterogeneity in treatment effects between durable and non-durable consumption documented in Section 4, we estimate welfare effects separately for these two retail product groups.

The estimated treatment effects give equal weight to all households in our endline data. To obtain welfare estimates that are representative at the village-level, we also re-estimate the treatment effects after weighting each household in our sample according to the fraction of the village population that resides within its sampling zone (inner or outer) in our endline data. These estimates are slightly smaller, but very similar (1.5 and 11 percentage points respectively), suggesting that our sampling procedure did not distort the average household in the village by much. For exposition, we report welfare estimates both with and without re-weighting households.

For the final set of required parameters in (8), the σ_g, we use the closest existing estimate of consumer demand across retailer choices in an emerging market context from recent work by

accounting for this additional effect.

Atkin et al. (in press) in Mexico. In particular, we use demand parameter estimates for households in Mexico with incomes comparable to those of rural Chinese households in our survey, which gives us $\sigma_N = 3.87$ for non-durables consumption and $\sigma_D = 3.85$ for durables consumption as baseline parameter values.[37]

To obtain standard errors for the welfare evaluation, we take into account that the treatment effects on ex-post e-commerce consumption shares are point estimates, not actual data points. We bootstrap the computation of expression 8 across 1000 iterations with random household re-sampling. Each iteration uses the mean and standard deviation of the estimated treatment effect on terminal share of retail consumption for durables and non-durables, and for each of the two household groups discussed above, and draws from a normal distribution around the mean of the respective point estimate of the treatment effects.

Table 11 reports the estimation results. The average reduction in retail cost of living among households who experienced the lifting of both logistical and transactional barriers is 0.81 percent. This effect increases to 5.5 percent among the roughly 14 percent of households who ever used the terminal for purchases. These effects are slightly lower at 0.71 and 4.8 percent respectively when weighting our sample households to represent the average population living in these villages. Underlying these effects are strong consumer gains in durable consumption: 2.9 percent for the average village household and 16.6 percent among users. For reference, retail consumption across all product groups accounts for on average 55 percent of total household expenditure among the rural households in the sample.

Distribution of the Gains from E-Commerce Integration

We now investigate the distribution of the gains from the arrival of e-commerce across households and villages. We use treatment effects from the heterogeneity specification in the last rows of Table 8, which includes all interactions with program treatment jointly estimated in one regression. We estimate this specification with the dependent variable being either household terminal share in durable or in non-durable retail consumption. For each sample household living in treatment villages without pre-existing parcel delivery, we compute a fitted value of the treatment effect on terminal retail consumption shares based on the primary earner's age, education, income per capita, residential distance to the planned terminal as well as distance to the nearest township center (remoteness).

We use these estimated effects for $\sum_{s \in S_g^C} \phi_{gsh}^{t1}$ in expression (8), and then plot the effect on household retail price indices flexibly across all sample households in treated villages. Figure 4 shows these plots for household income per capita (upper left), respondent age (upper right), distance to terminal (lower left) and distance to the nearest township center (lower right). These plots quantify the distribution of the gains to the average household, without restricting attention to users. The confidence intervals in these figures are based on sampling variation in household characteristics on the x-axis after clustering standard errors at the village-level.

The income plot shows that households in the 5th percentile of the income distribution on average experience a 0.25 percent reduction in retail cost of living due to the arrival of the new

[37] Atkin et al. (in press) estimate these parameters separately for richer and poorer households, and for food and non-food product groups. The parameters σ_N and σ_D that we use as our baseline refer to food and non-food product groups estimated among the poorer Mexican households respectively.

e-commerce option, which roughly quadruples to more than 1 percent for households at the 95th income percentile. A household with a 20 year-old primary earner on average experiences a reduction in retail cost of living of close to 2 percent, which drops below 1 percent past the age of 40. The gains are close to 1.5 percent on average in close residential proximity to the terminal and decrease to on average less than half a percent toward the largest distances in the sample. In contrast, villages in close proximity to the nearest township center experience a small reduction in retail cost of living that more than quadruple as distance from the nearest township reaches its maximum within our sample.

Overall, these figures reflect the significant heterogeneity in the program's impact on household consumption that we report in Table 8. We find that the benefits of e-commerce disproportionately accrue to households that are richer and younger, living in closer proximity to the e-commerce terminal within the villages, and in villages thar are relatively more remote.

Quantification Across Alternative Parameter Values

To account for uncertainty in the demand parameters, we compute results across alternative values of σ_N and σ_D, relative to our baseline parameterization ($\sigma_N = 3.87$ and $\sigma_D = 3.85$). In particular, we allow for household shopping demand to be either more or less price elastic across retailer options. Intuitively, the less price sensitive households are across retailers (i.e. the lower σ_g), the higher will be the implied consumer gains that are consistent with the observed household substitution to the new shopping option. We report results for a low-elasticity scenario with $\sigma_N = 2.87$ and $\sigma_D = 2.85$, and conversely for a high-elasticity scenario with $\sigma_N = 4.87$ and $\sigma_D = 4.85$. A priori, it is unclear which scenario is more likely in our current empirical context, relative to the baseline parameters estimated in Atkin et al. (in press) for similarly poor Mexican households living in urban areas. Rural Chinese households may be less sensitive to effective price differences across retailers due to higher shopping travel costs to a nearby town compared to urban Mexicans. Conversely, rural Chinese households may be intrinsically more price sensitive than Mexicans with similar real incomes.

Table A.10 reports the estimation results. As discussed above, assuming that rural Chinese shopping demand is less price elastic across retailer options than in our preferred parametrization yields significantly larger estimated welfare gains in retail consumption: a 1.25 percent reduction in cost of living for the average household in our sample and a 8.5 percent reduction for users. Conversely, assuming more price elastic shopping demand yields slightly smaller welfare effects of 0.6 and 4 percent respectively. For reference, the baseline estimates are 0.81 and 5.5 percent respectively.

Summary of Results

We find that the program leads to sizable gains in real income among households who are induced to ever use the e-commerce terminal. The welfare gains for the average rural household are more muted, suggesting strong heterogeneity in the effect of the arrival of e-commerce rather than broad-based welfare gains. The beneficiaries are on average richer and younger, live in closer proximity to the e-commerce terminals, and in villages that are farther away from the nearest township center. The welfare gains are driven by a significant reduction in household cost of living due to access to the new e-commerce shopping option that provides greater product variety, cheaper prices and a reduction in travel costs. These gains are strongest for durable product groups such as electronics and appliances.

7 Conclusion

The rapid growth of a number of rural e-commerce production hubs in China features prominently in recent policy reports and the popular press, and has attracted widespread policy interest. In this context, the Chinese government has launched the first nationwide e-commerce expansion program to remove the barriers to e-commerce development outside of cities. As the internet has become widely available in the countryside, this program aims to invest in removing the two main remaining barriers to e-commerce: the lack of modern transport logistics necessary for commercial parcel delivery and pickup (the logistical barrier), and the transitioning to non-traditional online user interfaces and paperless payments (the transactional barrier).

This paper uses this empirical context to study the economic consequences of e-commerce integration on the local economy, the underlying channels, and the distribution of the gains from e-commerce across households and villages. To this end, we combine an RCT that we implement across villages in collaboration with a large Chinese e-commerce firm with a new collection of microdata on household consumption, production and retail prices in the Chinese countryside.

The analysis provides several insights. We find that the program leads to sizable gains in real incomes among rural households who are induced to use the e-commerce terminal. These users represent about 14 percent of the rural household sample and 13 percent of the village population after adjusting for sampling weights. For the average rural household, including non-users, these gains are statistically significant but more muted. Underlying these effects, we find strong heterogeneity across households and villages. The beneficiaries of e-commerce are on average significantly younger, richer, live in closer proximity to the e-commerce terminal and in villages that are relatively more remote. Conditional on these characteristics, we do not find evidence that household education or the characteristics of the terminal managers affect the extent of household gains from e-commerce.

In terms of channels, we find significantly stronger economic gains among villages that were not previously serviced by commercial parcel delivery, suggesting that the program's gains are mainly due to overcoming the logistical, rather than transactional barrier. On the consumption side, we find that the e-commerce terminals offer lower prices, higher convenience and increased product variety compared to pre-existing local retail choices, both within the village and in nearby towns. The gains in household purchasing power are strongest for durable product groups, such as electronics and appliances. We also find suggestive evidence that the program led to additional product variety in pre-existing local stores, as their managers source new products through e-commerce. We find no evidence of significant pro-competitive effects on local retailer prices, on the other hand. On the production side, we find no evidence for significant effects on the local economy: online selling activity, purchases of production inputs, household incomes and entrepreneurship are not significantly affected by the arrival of the program. Overall, we find that the gains from e-commerce are driven by a reduction in local household cost of living that is mainly due to the direct gains from access to the new e-commerce shopping option for local households.

Using the firm's administrative database, we find little evidence on the consumption side suggesting that the adjustment to e-commerce takes longer than one year: the consumption-side uptake materializes within 2-4 months of entry and then remain mostly constant over time. On the production side, we find evidence that village-level out-shipments significantly increase over time beyond the 12-month window. However, the effect on total out-shipments remains relatively minor after more than two years post program entry, with a small upper-bound effect on local household incomes. Related to this, we do not find evidence that the survey data fails to pick up highly successful but rare tail events within villages on the production side that could in principle shift the mean effect on local household outcomes.

Overall, our findings suggest that e-commerce trading access offers significant economic gains to certain groups of the rural population, rather than being broad-based. Compared to the recent case studies highlighting a set of highly successful rural e-commerce production hubs, our analysis reveals that these prominent success stories are quite particular and not representative more generally. In the absence of complementary interventions, such as for example business training, access to credit, help standardizing production, or targeted online promotions, large and significant production-side effects appear unlikely to materialize for the average rural market place in the short to medium-run. In this light, future work aimed at better understanding the factors under which the arrival of e-commerce can have transformative impacts on the production side of the rural economy seems a promising agenda for future research in this area.

References

Akerman, A., Gaarder, I., & Mogstad, M. (2015). The skill complementarity of broadband internet. *The Quarterly Journal of Economics*, *130*(4), 1781–1824.

AliResearch. (2017). China's taobao villages. *Research Report*, 1741–1779.

Anderson, S. P., De Palma, A., & Thisse, J. F. (1992). *Discrete choice theory of product differentiation*. MIT press.

Atkin, D., Faber, B., & Gonzalez-Navarro, M. (in press). Retail globalization and household welfare: Evidence from mexico. *Journal of Political Economy*.

Banerjee, A., Duflo, E., & Qian, N. (2012). *On the road: Access to transportation infrastructure and economic growth in china* (Tech. Rep.). National Bureau of Economic Research.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2016). Roads, railroads and decentralization of Chinese cities. *Forthcoming, Review of Economics and Statistics*.

Bhuller, M., Havnes, T., Leuven, E., & Mogstad, M. (2013). Broadband Internet: An Information Superhighway to Sex Crime? *Review of Economic Studies*, *8*(4), 1237-1266.

Brynjolfsson, E., Hu, Y., & Smith, M. D. (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science*, *49*(11), 1580–1596.

Campante, F., Durante, R., & Sobbrio, F. (2013). Politics 2.0: The multifaceted effect of broadband internet on political participation. *Journal of the European Economic Association*.

Couture, V. (2016). Valuing the consumption benefits of urban density. *Working Paper, UC Berkeley*.

Disdier, A.-C., & Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics*, *90*(1), 37–48.

Donaldson, D. (in press). Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*.

Donaldson, D., & Hornbeck, R. (2016). Railroads and American economic growth: A market access approach. *The Quarterly Journal of Economics*, *131*(2), 799–858.

Einav, L., Klenow, P. J., Klopack, B., Levin, J. D., Levin, L., & Best, W. (2017). Assessing the gains from e-commerce. *Working paper, Stanford University*.

Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from China's national trunk highway system. *The Review of Economic Studies*, *81*(3), 1046–1070.

Falck, O., Gold, R., & Heblich, S. (2014). E-Lections: Voting Behavior and the Internet. *American Economic Review*, *107*(7), 2238-2265.

Fan, J., Tang, L., Zhu, W., & Zou, B. (2016). The alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce. *Michigan State University mimeo*.

Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *American Economic Review*, *84*(1), 155–57.

Forman, C., Goldfarb, A., & Greenstein, S. (2012). The Internet and Local Wages: A Puzzle. *American Economic Review*, *102*(1), 556-75.

FRED. (2016). E-commerce retail sales as a percent of total sales (ecompctsa). *Economic Data, Federal Reserve of St. Louis*.

Goldmanis, M., Hortaçsu, A., Syverson, C., & Emre, Ö. (2010). E-commerce and the market structure of retail industries. *The Economic Journal*, *120*(545), 651–682.

Goyal, A. (2010). Information, direct access to farmers, and rural market performance in central india. *American Economic Journal: Applied Economics*, *2*(3), 22–45.

Hamory, J., Kleemans, M., Li, N., & Miguel, E. (2016). Individual ability and selection into migration in Kenya. *Mimeo, UC Berkeley*.

Handbury, J. (2013). Are poor cities cheap for everyone? Non-homotheticity and the cost of living across us cities. *Working Paper, Wharton*.

Handbury, J., & Weinstein, D. E. (2015). Goods prices and availability in cities. *The Review of Economic Studies*, *82*(1), 258–296.

Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The economics of new goods* (pp. 207–248). University of Chicago Press.

Hausman, J. A., & Leonard, G. K. (2002). The competitive effects of a new product introduction: A case study. *Journal of Industrial Economics*, *50*(3), 237–263.

Head, K., & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics* (Vol. 4, pp. 131–195). Elsevier.

Hicks, J. R. (1940). The valuation of the social income. *Economica*, 105–124.

Hjort, J., & Poulsen, J. (2017). The arrival of fast internet and skilled job creation in Africa. *Columbia University Working Paper*.

Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector. *The quarterly journal of economics*, *122*(3), 879–924.

Lagakos, D., Mobarak, M., & Waugh, M. (2016). Urban-rural wage gaps in developing countries: spatial misallocation or efficient sorting. *Mimeo, UC San Diego, Yale, and NYU*.

MCIT. (2016). Egypt's national e-commerce strategy. *Ministry of Communications and Information Technology Report*.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more t in experiments. *Journal of Development Economics*, *99*(2), 210–221.

McKinsey. (2016). China's e-tail revolution. *Report, McKinsey Global Institute*.

MEITY. (2016). Digital india: Driving e-commerce in rural and semi urban India. *Ministry of Electronics and IT Report*.

Miguel, E., & Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, *72*(1), 159–217.

PFSweb. (2016). China's e-commerce market 2015. *Annual Market Report*.

PM. (2016). Vietnam's e-commerce development masterplan. *Vietnam's Office of the Prime Minister Report*.

Statista. (2016). Online retail statistics for china. *Market Statistics*.

Topalova, P. (2010). Factor immobility and regional impacts of trade liberalization: Evidence on poverty from India. *American Economic Journal: Applied Economics*, *2*(4), 1–41.

UNCTAD. (2016a). E-commerce opens new opportunities for developing countries. *UNCTAD/PRESS/U14/IN/2016/011*.

UNCTAD. (2016b). Unctad e-commerce index 2016. *Report, United Nations Conference on Trade and Development, Geneva*.

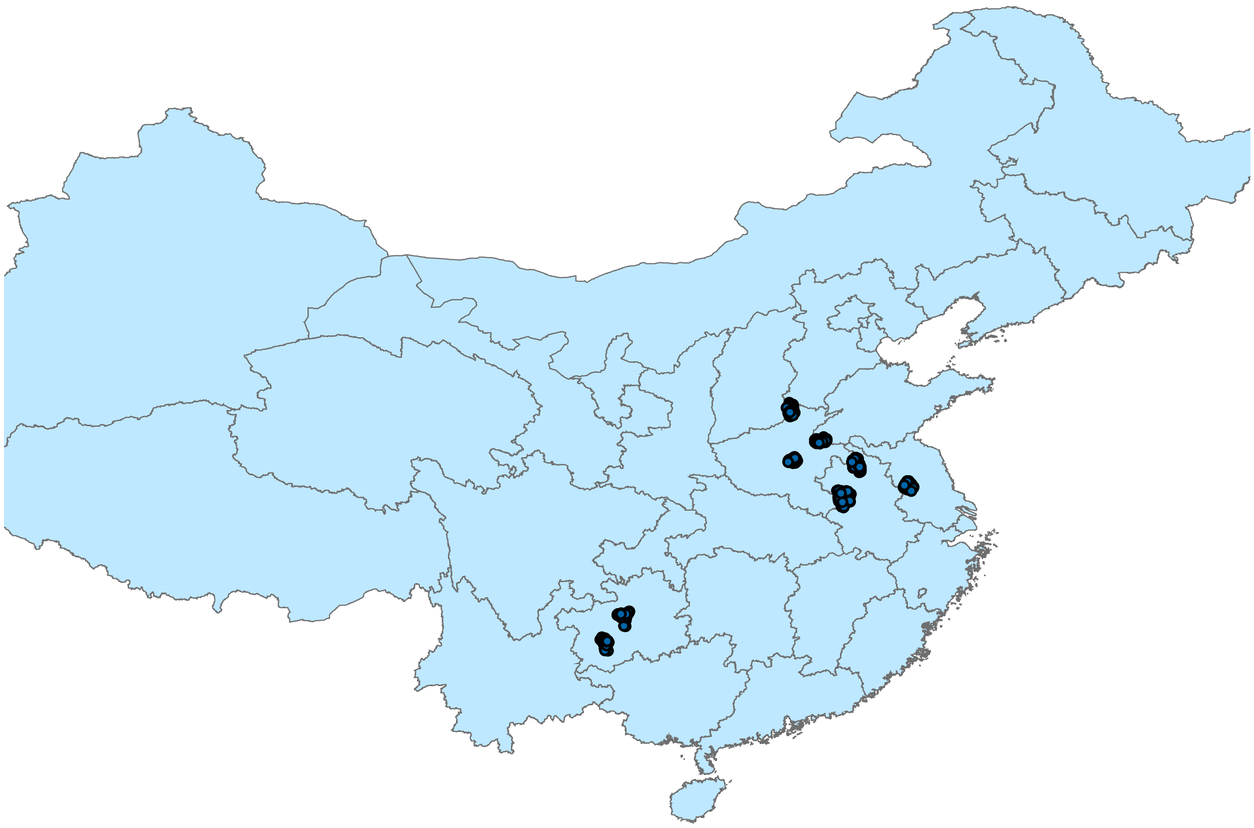
WTO. (2013). E-commerce in developing countries: Opportunities and challenges for small and medium-sized enterprises. *Policy Report, World Trade Organization, Geneva*.

Young, A. (2013). Inequality, the urban-rural gap and migration. *The Quarterly Journal of Economics*, 1727–1785.

8 Figures and Tables

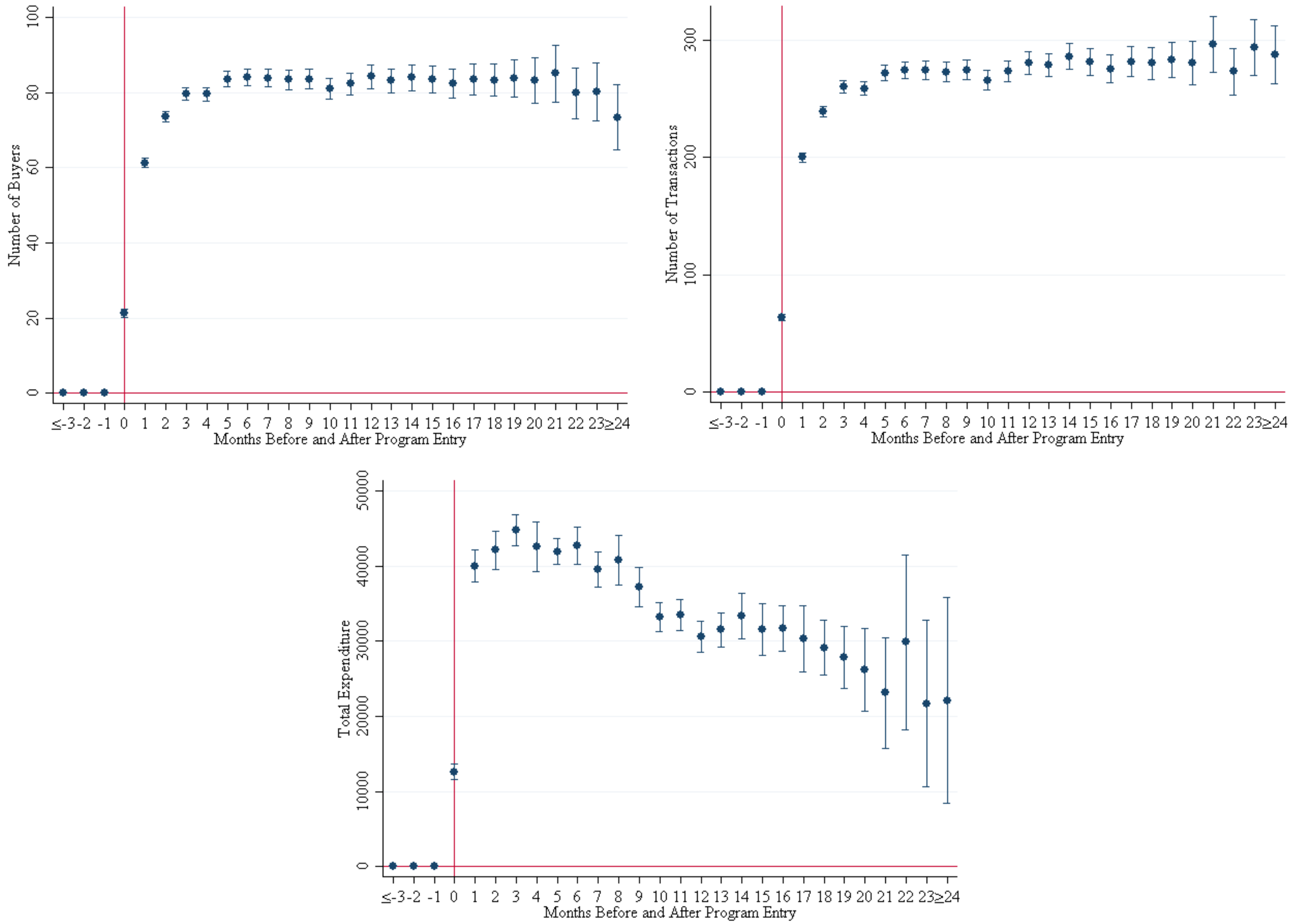
8.1 Figures

Figure 1: Provinces and Counties Where RCT Was Implemented



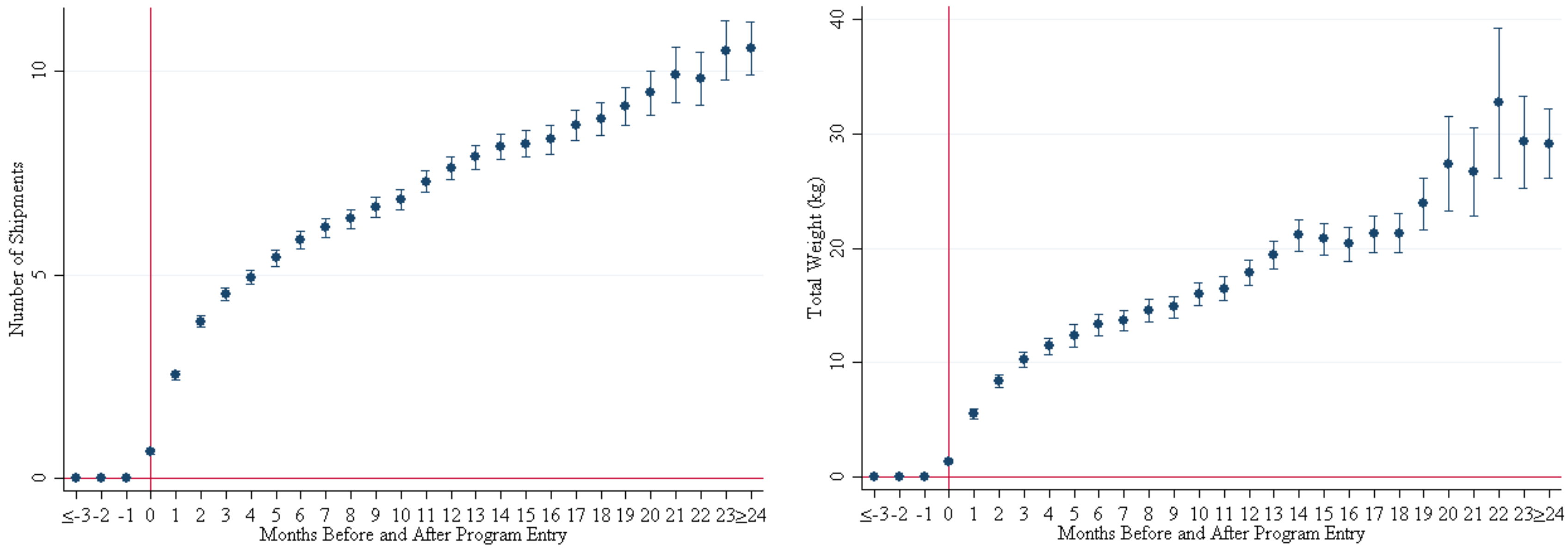
Notes: The boundaries indicate Mainland Chinese provinces. The dots indicate participating villages in the 8 counties where the RCT takes place. See Section 2 for discussion.

Figure 2: Timeline of Adjustment: Consumption (Terminal-Level)



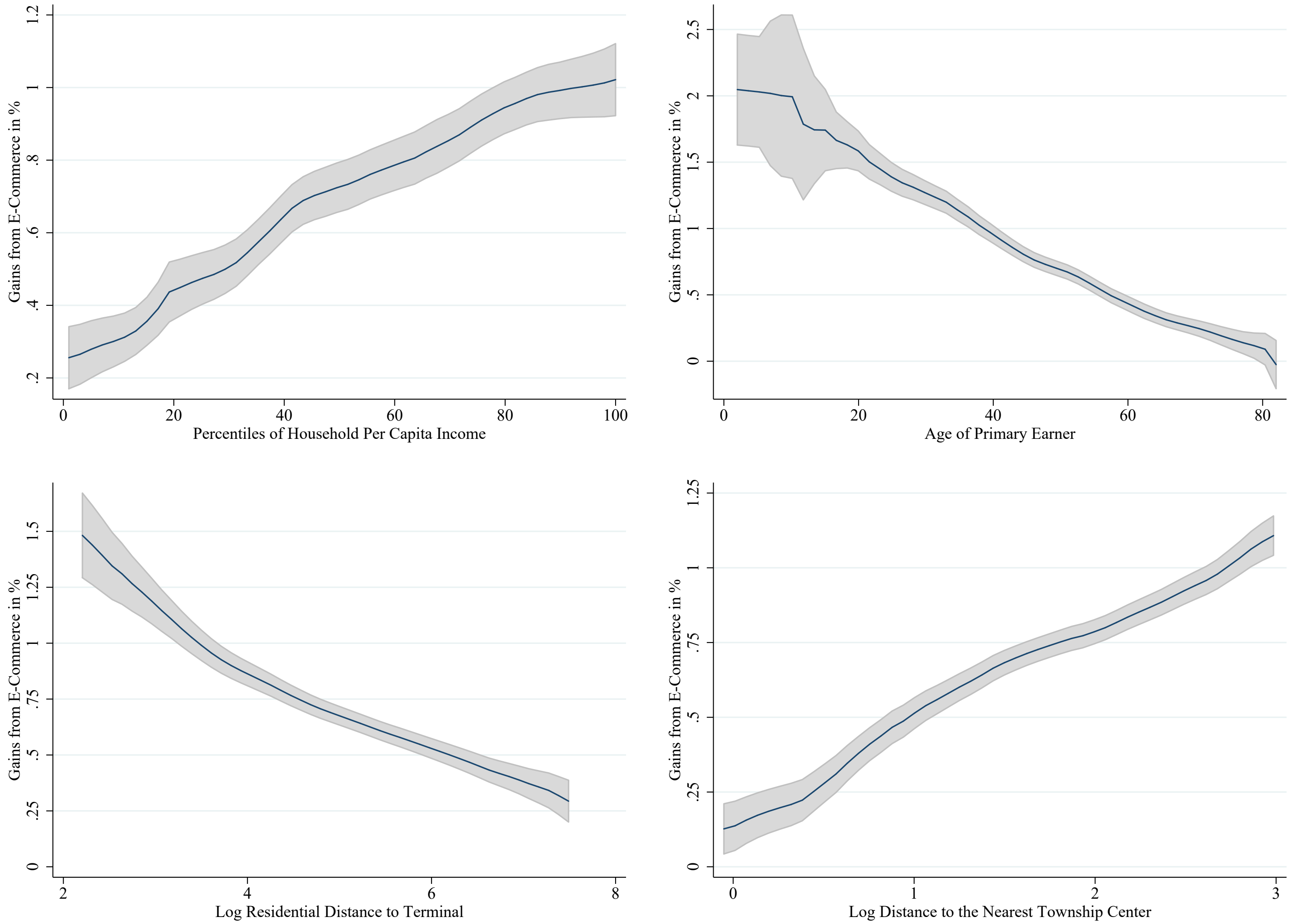
Notes: See Section 5 for discussion. Standard errors are clustered at the level of village terminals. * 10%, ** 5%, *** 1% significance levels.

Figure 3: Timeline of Adjustment: Selling (Terminal-Level)



Notes: See Section 5 for discussion. Standard errors are clustered at the level of village terminals. * 10%, ** 5%, *** 1% significance levels.

Figure 4: Heterogeneity of Welfare Effect



Notes: See Section 6 for discussion. Gains are expressed in terms of percentage point reductions of retail cost of living. Confidence intervals are based on standard errors clustered at the level of villages.

8.2 Tables

Table 1: Survey Data Statistics

		Full Sample at Baseline	Treatment Villages at Baseline	Control Villages at Baseline	P-Value (Treat-Control=0)	Control Villages at Endline
Panel A: Individual Level						
Age	Median	44.000	44.000	43.000		46.000
	Mean	38.950	39.329	38.407	0.208	39.943
	Standard Deviation	23.580	23.658	23.460		23.759
	Number of Obs	8491	5001	3490		4194
Gender (Female=1)	Median	1.000	1.000	1.000		1.000
	Mean	0.534	0.526	0.546	0.025	0.537
	Standard Deviation	0.499	0.499	0.498		0.499
	Number of Obs	8484	5001	3483		4188
Employed (for age>15) (Yes=1)	Median	1.000	1.000	1.000		1.000
	Mean	0.767	0.766	0.769	0.882	0.762
	Standard Deviation	0.423	0.424	0.422		0.426
	Number of Obs	6070	3590	2480		3015
Peasant (for age>15) (Yes=1)	Median	1.000	1.000	1.000		1.000
	Mean	0.527	0.527	0.526	0.971	0.513
	Standard Deviation	0.499	0.499	0.499		0.500
	Number of Obs	6369	3760	2609		3144
No Schooling (for age>15) (No School=1)	Median	0.000	0.000	0.000		0.000
	Mean	0.270	0.273	0.266	0.745	0.319
	Standard Deviation	0.444	0.446	0.442		0.466
	Number of Obs	6368	3758	2610		3132
Panel B: Household Level						
Household Size	Median	3.000	3.000	3.000		3.00
	Mean	3.114	3.053	3.205	0.075	2.987
	Standard Deviation	1.422	1.420	1.421		1.40
	Number of Obs	2740	1647	1093		1405
Gender of Primary Earner (Female=1)	Median	0.000	0.000	0.000		0.00
	Mean	0.288	0.295	0.276	0.457	0.295
	Standard Deviation	0.453	0.456	0.447		0.46
	Number of Obs	2547	1530	1017		1348
Primary Earner Self-Employed (Yes=1)	Median	0.000	0.000	0.000		0.00
	Mean	0.073	0.087	0.053	0.036	0.072
	Standard Deviation	0.261	0.282	0.224		0.26
	Number of Obs	2549	1531	1018		1348
Primary Earner Is Peasant (Yes=1)	Median	1.000	1.000	1.000		1.00
	Mean	0.590	0.600	0.577	0.620	0.587
	Standard Deviation	0.492	0.490	0.494		0.49
	Number of Obs	2549	1531	1018		1348
Household Monthly Income Per Capita in RMB	Median	350.000	339.000	375.000		466.67
	Mean	876.412	841.198	929.473	0.365	1028.960
	Standard Deviation	1717.456	1687.169	1761.560		2005.31
	Number of Obs	2740	1647	1093		1405
Household Monthly Retail Expenditure Per Capita in RMB	Median	381.000	372.833	400.500		364.00
	Mean	732.017	663.034	835.966	0.135	686.616
	Standard Deviation	2304.540	1139.788	3368.220		1512.06
	Number of Obs	2735	1644	1091		1405

Notes: See Section 2 for discussion.

Table 2: Survey Data Statistics (Continued)

		Full Sample at Baseline	Treatment Villages at Baseline	Control Villages at Baseline	P-Value (Treat-Control=0)	Control Villages at Endline
Panel B: Household Level (Continued)						
Share of Retail Expenditure Outside of Village	Median	0.553	0.489	0.623		0.60
	Mean	0.500	0.470	0.545	0.193	0.531
	Standard Deviation	0.395	0.402	0.379		0.38
	Number of Obs	2720	1637	1083		1397
Distance in Meters to Planned Terminal Location	Median	231.556	232.891	231.454		203.63
	Mean	290.346	293.364	285.797	0.789	286.631
	Standard Deviation	243.450	247.778	236.820		267.06
	Number of Obs	2740	1647	1093		1405
Any Member of the Household Has Ever Used the Internet (Yes=1)	Median	0.000	0.000	0.000		0.00
	Mean	0.368	0.354	0.390	0.249	0.427
	Standard Deviation	0.482	0.478	0.488		0.49
	Number of Obs	2739	1646	1093		1402
Household Owns a Smartphone (Yes=1)	Median	1.000	1.000	1.000		1.00
	Mean	0.526	0.509	0.552	0.153	0.551
	Standard Deviation	0.499	0.500	0.498		0.50
	Number of Obs	2731	1642	1089		1400
Share of Household Monthly Expenditure on E-Commerce Deliveries	Median	0.000	0.000	0.000		0.00
	Mean	0.007	0.006	0.007	0.693	0.008
	Standard Deviation	0.050	0.046	0.057		0.05
	Number of Obs	2720	1637	1083		1397
Share of E-Commerce Sales in Household Monthly Income	Median	0.000	0.000	0.000		0.00
	Mean	0.003	0.001	0.006	0.103	0.003
	Standard Deviation	0.052	0.030	0.074		0.05
	Number of Obs	2055	1244	811		1161
Panel C: Local Retail Survey						
Number of Stores at Village Level	Median	3.00	3.00	2.00		2.00
	Mean	4.15	4.38	3.79	0.33	3.61
	Standard Deviation	2.94	2.91	2.98		2.99
	Number of Obs	99	60	39		38
Establishment Space in Square Meters	Median	50.00	50.00	40.00		50.00
	Mean	99.07	74.42	146.76	0.35	121.33
	Standard Deviation	320.38	89.60	532.73		375.35
	Number of Obs	361	238	123		126
Number of Establishment's New Products Added Over Last Month	Median	0.00	0.00	0.00		0.00
	Mean	1.43	1.56	1.17	0.57	0.63
	Standard Deviation	7.44	8.88	3.42		2.26
	Number of Obs	330	215	115		126
Prices of All Retail Consumption (9 Product Groups) in RMB	Median	7.00	7.00	6.00		6.00
	Mean	71.03	76.74	61.43	0.47	71.23
	Standard Deviation	411.24	433.67	370.33		390.31
	Number of Obs	9382	5884	3498		3259
Prices of Business or Production Input in RMB	Median	10.00	10.00	8.80		9.00
	Mean	45.63	42.88	49.78	0.76	43.84
	Standard Deviation	195.09	206.23	177.46		97.92
	Number of Obs	444	267	177		111

Notes: See Section 2 for discussion.

Table 3: Firm's Transaction Data

	Number of Purchase Transactions	Number of Buyers	Number of Out-Shipments	Number of Terminals	Number of Counties	Number of Provinces	Number of Days	Number of Months	Sum of Payments (RMB)	Sum of Out-Shipments (Weight in kg)
Full Sample	27,270,532	3,785,019	500,743	11,941	175	5	547	18	4,480,424,896	1,169,673
3 Provinces	20,647,373	2,832,872	442,319	8,561	116	3	547	18	3,409,227,245	1,019,373
8 Counties	1,835,897	216,529	44,148	706	8	3	503	17	330,930,097	95,908
RCT Villages	130,769	15,099	3,158	40	8	3	482	16	17,618,900	7,817

Notes: The table provides information from the purchase and the sales transaction databases. The purchasing database covers all village transaction in 5 provinces over the period November 2015 until April 2017. The sales transaction database covers all out-shipments from the same locations over the period January 2016 to April 2017. See Section 2 for discussion.

Table 4: Average Effects: Consumption

Dependent Variables		Intent to Treat	Treatment on Treated	Log Distance (IV using Treat)	Dependent Variables		Intent to Treat	Treatment on Treated	Log Distance (IV using Treat)
Monthly Total Retail Expenditure Per Capita	Treat or Log Dist	-21.93	-40.92	11.15	Share of Terminal in Monthly Tobacco and Alcohol (2)	Treat or Log Dist	0.000608	0.00123	-0.000352
		(31.96)	(60.19)	(16.29)			(0.000515)	(0.00109)	(0.000306)
	R-Squared	0.038				R-Squared	0.001		
	First Stage F-Stat		43.92	42.45		First Stage F-Stat		33.02	27.08
	Number of Obs	3,434	3,434	3,434		Number of Obs	1,653	1,653	1,653
Household Has Ever Bought Something at Terminal (Yes=1)	Treat or Log Dist	0.0480***	0.0886***	-0.0241***	Share of Terminal in Monthly Medicine and Health Products (3)	Treat or Log Dist	0.000693	0.00126	-0.000344
		(0.0169)	(0.0271)	(0.00721)			(0.000689)	(0.00124)	(0.000339)
	R-Squared	0.008				R-Squared	0.000		
	First Stage F-Stat		45.56	43.80		First Stage F-Stat		51.06	46.74
	Number of Obs	3,518	3,518	3,518		Number of Obs	2,416	2,416	2,416
Household Has Bought Something at Terminal in Past Month (Yes=1)	Treat or Log Dist	0.0263***	0.0490***	-0.0134***	Share of Terminal in Monthly Clothing and Accessories (4)	Treat or Log Dist	0.0465***	0.0734***	-0.0205***
		(0.00981)	(0.0171)	(0.00458)			(0.0140)	(0.0216)	(0.00603)
	R-Squared	0.009				R-Squared	0.019		
	First Stage F-Stat		43.93	42.23		First Stage F-Stat		70.69	56.57
	Number of Obs	3,482	3,482	3,482		Number of Obs	1,269	1,269	1,269
Share of Online Terminal in Total Monthly Retail Expenditure	Treat or Log Dist	0.00666***	0.0124***	-0.00338***	Share of Terminal in Monthly Other Household Products (5)	Treat or Log Dist	0.00430	0.00804	-0.00225
		(0.00239)	(0.00434)	(0.00117)			(0.00395)	(0.00713)	(0.00198)
	R-Squared	0.006				R-Squared	0.001		
	First Stage F-Stat		44.03	42.34		First Stage F-Stat		43.87	39.89
	Number of Obs	3,434	3,434	3,434		Number of Obs	2,336	2,336	2,336
Share of Terminal in Monthly Business Inputs	Treat or Log Dist	-0.00715	-0.0154	0.00433	Share of Terminal in Monthly Heating, Fuel and Gas (6)	Treat or Log Dist	0	0	0
		(0.00778)	(0.0191)	(0.00545)			(0)	(0)	(0)
	R-Squared	0.003				R-Squared	.		
	First Stage F-Stat		16.46	14.96		First Stage F-Stat		.	.
	Number of Obs	1,207	1,207	1,207		Number of Obs	1,463	1,463	1,463
Share of Terminal in Monthly Non-Durables	Treat or Log Dist	0.00536***	0.00999***	-0.00272***	Share of Terminal in Monthly Furniture and Appliances (7)	Treat or Log Dist	0.0546**	0.0908**	-0.0248**
		(0.00195)	(0.00355)	(0.000956)			(0.0217)	(0.0368)	(0.00989)
	R-Squared	0.003				R-Squared	0.019		
	First Stage F-Stat		44.11	42.33		First Stage F-Stat		47.51	44.31
	Number of Obs	3,433	3,433	3,433		Number of Obs	380	380	380
Share of Terminal in Monthly Durables	Treat or Log Dist	0.0398**	0.0669**	-0.0188**	Share of Terminal in Monthly Electronics (8)	Treat or Log Dist	0.0697**	0.110**	-0.0322**
		(0.0159)	(0.0261)	(0.00736)			(0.0345)	(0.0522)	(0.0152)
	R-Squared	0.011				R-Squared	0.024		
	First Stage F-Stat		52.64	41.27		First Stage F-Stat		43.20	26.28
	Number of Obs	768	768	768		Number of Obs	232	232	232
Share of Terminal in Monthly Food and Beverages (1)	Treat or Log Dist	0.00121	0.00223	-0.000606	Share of Terminal in Monthly Transport Equipment (9)	Treat or Log Dist	0.0353*	0.0554*	-0.0162*
		(0.000823)	(0.00152)	(0.000414)			(0.0201)	(0.0313)	(0.00935)
	R-Squared	0.001				R-Squared	0.014		
	First Stage F-Stat		45.63	43.70		First Stage F-Stat		43.07	31.48
	Number of Obs	3,359	3,359	3,359		Number of Obs	141	141	141

Notes: See Section 4 for discussion. Standard errors are clustered at the level of villages. * 10%, ** 5%, *** 1% significance levels.

Table 5: Average Effects: Incomes

Dependent Variables		Intent to Treat	Treatment on Treated	Log Distance (IV using Treat)	Dependent Variables		Intent to Treat	Treatment on Treated	Log Distance (IV using Treat)
Monthly Income Per Capita in RMB	Treat or Log Dist	-7.838	-14.48	3.974	Member of Household Has Ever Sold Online (Yes=1)	Treat or Log Dist	-0.00700	-0.0129	0.00353
		(70.78)	(129.9)	(35.61)			(0.00562)	(0.0104)	(0.00282)
	R-Squared	0.038				R-Squared	0.347		
	First Stage F-Stat		45.33	42.83		First Stage F-Stat		45.30	42.71
	Number of Obs	3,437	3,437	3,437		Number of Obs	3,504	3,504	3,504
Monthly Income Per Capita Net of Costs in RMB	Treat or Log Dist	-20.09	-37.20	10.19	Member of Household Has Sold Online In Past Month (Yes=1)	Treat or Log Dist	-0.00132	-0.00244	0.000667
		(70.80)	(129.9)	(35.51)			(0.00237)	(0.00438)	(0.00119)
	R-Squared	0.037				R-Squared	0.038		
	First Stage F-Stat		44.78	42.54		First Stage F-Stat		44.30	42.34
	Number of Obs	3,390	3,390	3,390		Number of Obs	3,498	3,498	3,498
Monthly Income Per Capita Net of Transfers in RMB	Treat or Log Dist	-12.55	-23.21	6.360	Online Sales in Past Month in RMB	Treat or Log Dist	-10.09	-18.75	5.109
		(72.18)	(132.4)	(36.25)			(12.89)	(23.94)	(6.504)
	R-Squared	0.051				R-Squared	0.012		
	First Stage F-Stat		45.16	42.67		First Stage F-Stat		44.26	42.39
	Number of Obs	3,445	3,445	3,445		Number of Obs	3,498	3,498	3,498
Annual Income Per Capita in RMB	Treat or Log Dist	-45.95	-85.08	23.33	Share of Online Sales in Household Monthly Income	Treat or Log Dist	-0.00120	-0.00224	0.000614
		(586.9)	(1,080)	(296.3)			(0.00176)	(0.00330)	(0.000901)
	R-Squared	0.046				R-Squared	0.032		
	First Stage F-Stat		44.77	42.23		First Stage F-Stat		41.62	38.41
	Number of Obs	3,388	3,388	3,388		Number of Obs	2,830	2,830	2,830
Monthly Agricultural Income Per Capita	Treat or Log Dist	-70.23	-130.3	35.61	Primary Earner Working As Peasant (Yes=1)	Treat or Log Dist	-0.0229	-0.0425	0.0116
		(140.3)	(257.7)	(70.34)			(0.0319)	(0.0597)	(0.0164)
	R-Squared	0.033				R-Squared	0.140		
	First Stage F-Stat		44.23	42.33		First Stage F-Stat		44.42	41.58
	Number of Obs	3,448	3,448	3,448		Number of Obs	3,327	3,327	3,327
Monthly Non-Agricultural Income Per Capita	Treat or Log Dist	-46.65	-86.06	23.55	Member of Household Started a Business Over Last 6 Months (Yes=1)	Treat or Log Dist	-0.00802	-0.0149	0.00407
		(137.3)	(249.6)	(68.28)			(0.00631)	(0.0120)	(0.00327)
	R-Squared	0.157				R-Squared	0.001		
	First Stage F-Stat		45.74	43.51		First Stage F-Stat		44.37	42.34
	Number of Obs	3,441	3,441	3,441		Number of Obs	3,468	3,468	3,468
Weekly Hours Worked by Primary Earner	Treat or Log Dist	1.008	1.879	-0.516	New Business Selling in Part Online (Yes=1)	Treat or Log Dist	0.000212	0.000394	-0.000108
		(3.383)	(6.285)	(1.723)			(0.00159)	(0.00294)	(0.000803)
	R-Squared	0.000				R-Squared	0.000		
	First Stage F-Stat		43.80	41.21		First Stage F-Stat		44.33	42.37
	Number of Obs	3,310	3,310	3,310		Number of Obs	3,468	3,468	3,468
Weekly Hours Worked by Secondary Earner	Treat or Log Dist	-0.0606	-0.110	0.0317					
		(3.886)	(7.002)	(2.020)					
	R-Squared	0.000							
	First Stage F-Stat		45.39	40.21					
	Number of Obs	1,866	1,866	1,866					

Notes: See Section 4 for discussion. Standard errors are clustered at the level of villages. * 10%, ** 5%, *** 1% significance levels.

Table 6: Average Effects: Local Retail Prices

Dependent Variables		Intent to Treat	Treatment on Treated	Dependent Variables		Intent to Treat	Treatment on Treated
Log Prices (All)	Treat	0.0189	0.0352	Log Prices of Food and Beverages (1)	Treat	0.0368**	0.0706*
		(0.0142)	(0.0263)			(0.0185)	(0.0375)
	R-Squared	0.893	0.893		R-Squared	0.870	0.870
	First Stage F-Stat		41.66		First Stage F-Stat		39.37
	Number of Obs	6,877	6,877		Number of Obs	3,686	3,686
Product Replacement Dummy (Not Counting Store Closures) (Yes=1)	Treat	-0.00516	-0.00983	Log Prices of Tobacco and Alcohol (2)	Treat	0.0212	0.0421
		(0.00947)	(0.0181)			(0.0340)	(0.0662)
	R-Squared	0.000	-0.002		R-Squared	0.809	0.810
	First Stage F-Stat		39.82		First Stage F-Stat		32.39
	Number of Obs	8,956	8,956		Number of Obs	1,071	1,071
Store Closure (at Product Level) (Yes=1)	Treat	0.00124	0.00236	Log Prices of Medicine and Health Products (3)	Treat	-0.0474	-0.0756
		(0.0294)	(0.0556)			(0.0741)	(0.122)
	R-Squared	0.000	0.000		R-Squared	0.794	0.795
	First Stage F-Stat		39.82		First Stage F-Stat		19.18
	Number of Obs	8,956	8,956		Number of Obs	266	266
Number of New Products Per Store	Treat	2.194**	4.020*	Log Prices of Clothing and Accessories (4)	Treat	0.0809	0.115
		(1.073)	(2.278)			(0.111)	(0.158)
	R-Squared	0.277	0.212		R-Squared	0.845	0.842
	First Stage F-Stat		19.69		First Stage F-Stat		42.80
	Number of Obs	312	312		Number of Obs	152	152
Store Owner Sources Products Online (Yes=1)	Treat	-0.00145	-0.00261	Log Prices of Other Household Products (5)	Treat	-0.0328	-0.0619
		(0.0258)	(0.0461)			(0.0382)	(0.0744)
	R-Squared	0.000	-0.001		R-Squared	0.756	0.755
	First Stage F-Stat		23.76		First Stage F-Stat		28.85
	Number of Obs	341	341		Number of Obs	1,268	1,268
Log Prices of Business Inputs	Treat	0.00229	0.00337	Log Prices of Heating, Fuel and Gas (6)	Treat	-0.0115	-0.0440
		(0.129)	(0.186)			(0.0955)	(0.332)
	R-Squared	0.811	0.811		R-Squared	0.007	-0.095
	First Stage F-Stat		24.86		First Stage F-Stat		0.795
	Number of Obs	237	237		Number of Obs	12	12
Log Prices of Non-Durables	Treat	0.0211	0.0398	Log Prices of Furniture and Appliances (7)	Treat	-0.0347	-0.0617
		(0.0146)	(0.0276)			(0.0881)	(0.156)
	R-Squared	0.860	0.860		R-Squared	0.952	0.953
	First Stage F-Stat		40.36		First Stage F-Stat		6.757
	Number of Obs	6,455	6,455		Number of Obs	109	109
Log Prices of Durables	Treat	-0.0320	-0.0522	Log Prices of Electronics (8)	Treat	-0.0892	-0.163
		(0.0711)	(0.115)			(0.305)	(0.570)
	R-Squared	0.951	0.952		R-Squared	0.884	0.890
	First Stage F-Stat		9.753		First Stage F-Stat		3.180
	Number of Obs	185	185		Number of Obs	23	23
				Log Prices of Transport Equipment (9)	Treat	0.0297	0.0398
						(0.0840)	(0.110)
					R-Squared	0.946	0.946
					First Stage F-Stat		22.67
					Number of Obs	53	53

Notes: See Section 4 for discussion. Standard errors are clustered at the level of villages. * 10%, ** 5%, *** 1% significance levels.

Table 7: Effects of Logistical vs Transactional Barriers

Effects on Consumption

Dept Variables		Intent to Treat	Treatment on the Treated	Log Distance (IV Using Treat)
Monthly Total Retail Expenditure Per Capita	Treat or Log Dist	-26.72	-49.03	13.55
		(36.25)	(67.96)	(18.65)
	Treat or Log Dist * Delivery	31.42	58.59	-15.88
		(69.33)	(140.5)	(35.96)
	First Stage F-Stat		2.388	2.466
	Number of Obs	3,434	3,434	3,434
Household Has Ever Bought Something at Terminal (Yes=1)	Treat or Log Dist	0.0573***	0.105***	-0.0289***
		(0.0190)	(0.0288)	(0.00776)
	Treat or Log Dist * Delivery	-0.0603**	-0.110**	0.0304***
		(0.0251)	(0.0438)	(0.0113)
	First Stage F-Stat		2.683	2.754
	Number of Obs	3,518	3,518	3,518
Household Has Bought Something at Terminal in Last Month (Yes=1)	Treat or Log Dist	0.0329***	0.0604***	-0.0167***
		(0.0111)	(0.0189)	(0.00518)
	Treat or Log Dist * Delivery	-0.0422***	-0.0790**	0.0214**
		(0.0155)	(0.0329)	(0.00855)
	First Stage F-Stat		2.513	2.577
	Number of Obs	3,482	3,482	3,482
Share of Expenditure at Terminal in Total Monthly Retail Expenditure	Treat or Log Dist	0.00796***	0.0146***	-0.00405***
		(0.00274)	(0.00488)	(0.00134)
	Treat or Log Dist * Delivery	-0.00833***	-0.0153***	0.00424***
		(0.00294)	(0.00542)	(0.00147)
	First Stage F-Stat		2.413	2.483
	Number of Obs	3,434	3,434	3,434
Share of Expenditure at Terminal in Total Monthly Business Inputs	Treat or Log Dist	-0.00830	-0.0190	0.00548
		(0.00827)	(0.0222)	(0.00656)
	Treat or Log Dist * Delivery	0.0158	0.0296	-0.00790
		(0.0105)	(0.0241)	(0.00685)
	First Stage F-Stat		6.346	5.536
	Number of Obs	1,207	1,207	1,207
Share of Expenditure at Terminal in Total Monthly Non-Durables	Treat or Log Dist	0.00637***	0.0117***	-0.00324***
		(0.00225)	(0.00400)	(0.00110)
	Treat or Log Dist * Delivery	-0.00646**	-0.0119***	0.00329***
		(0.00246)	(0.00452)	(0.00122)
	First Stage F-Stat		2.413	2.483
	Number of Obs	3,433	3,433	3,433
Share of Expenditure at Terminal in Total Monthly Durables	Treat or Log Dist	0.0486***	0.0807***	-0.0233***
		(0.0177)	(0.0284)	(0.00822)
	Treat or Log Dist * Delivery	-0.0694***	-0.118***	0.0324***
		(0.0258)	(0.0442)	(0.0121)
	First Stage F-Stat		3.150	17.74
	Number of Obs	768	768	768

Effects on Incomes

Dept Variables		Intent to Treat	Treatment on the Treated	Log Distance (IV Using Treat)
Monthly Income Per Capita in RMB	Treat or Log Dist	-14.99	-27.14	7.579
		(77.55)	(140.1)	(39.08)
	Treat or Log Dist * Delivery	50.29	97.16	-25.08
		(171.2)	(339.1)	(86.90)
	First Stage F-Stat		2.694	2.737
	Number of Obs	3,437	3,437	3,437
Monthly Income Per Capita Net of Costs in RMB	Treat or Log Dist	-20.24	-37.09	10.33
		(77.47)	(140.5)	(39.07)
	Treat or Log Dist * Delivery	6.011	9.303	-3.362
		(167.6)	(317.4)	(81.28)
	First Stage F-Stat		2.810	2.852
	Number of Obs	3,390	3,390	3,390
Monthly Income Per Capita Net of Transfers in RMB	Treat or Log Dist	-13.87	-25.27	7.041
		(77.86)	(140.7)	(39.18)
	Treat or Log Dist * Delivery	12.70	23.04	-6.473
		(188.3)	(367.2)	(93.22)
	First Stage F-Stat		2.635	2.696
	Number of Obs	3,445	3,445	3,445
Annual Income Per Capita in RMB	Treat or Log Dist	70.33	124.2	-34.68
		(645.0)	(1,168)	(325.6)
	Treat or Log Dist * Delivery	-734.1	-1,462	368.3
		(1,484)	(2,755)	(692.5)
	First Stage F-Stat		2.501	2.603
	Number of Obs	3,388	3,388	3,388
Member of Household Has Ever Sold Online (Yes=1)	Treat or Log Dist	-0.00857	-0.0156	0.00433
		(0.00608)	(0.0111)	(0.00309)
	Treat or Log Dist * Delivery	0.0102	0.0188	-0.00513
		(0.0141)	(0.0280)	(0.00715)
	First Stage F-Stat		2.561	2.598
	Number of Obs	3,504	3,504	3,504
Share of Online Sales in Household Monthly Income	Treat or Log Dist	-0.00172	-0.00316	0.000882
		(0.00210)	(0.00387)	(0.00108)
	Treat or Log Dist * Delivery	0.00282	0.00540	-0.00145
		(0.00233)	(0.00441)	(0.00121)
	First Stage F-Stat		2.402	2.342
	Number of Obs	2,830	2,830	2,830
Primary Earner Working as Peasant (Yes=1)	Treat or Log Dist	-0.0192	-0.0352	0.00979
		(0.0341)	(0.0624)	(0.0174)
	Treat or Log Dist * Delivery	-0.0284	-0.0609	0.0143
		(0.0813)	(0.185)	(0.0464)
	First Stage F-Stat		2.503	2.533
	Number of Obs	3,327	3,327	3,327
Member of Household Has Started a Business Over Last 6 Months (Yes=1)	Treat or Log Dist	-0.00328	-0.00601	0.00167
		(0.00635)	(0.0116)	(0.00322)
	Treat or Log Dist * Delivery	-0.0297	-0.0604	0.0149
		(0.0183)	(0.0536)	(0.0130)
	First Stage F-Stat		2.517	2.566
	Number of Obs	3,468	3,468	3,468

Effects on Retail Prices

Dept Variables		Intent to Treat	Treatment on the Treated
Log Prices (All)	Treat	0.0114	0.0215
		(0.0144)	(0.0273)
	Treat * Delivery	0.0417	0.0739
		(0.0377)	(0.0572)
	First Stage F-Stat		17.26
	Number of Obs	6,877	6,877
Product Replacement Dummy (Not Counting Store Closures) (Yes=1)	Treat	-0.00680	-0.0129
		(0.0108)	(0.0206)
	Treat * Delivery	0.00907	0.0173
		(0.0213)	(0.0415)
	First Stage F-Stat		2.648
	Number of Obs	8,956	8,956
Store Closure (at Product Level) (Yes=1)	Treat	0.00111	0.00209
		(0.0355)	(0.0668)
	Treat * Delivery	0.000779	0.00162
		(0.0423)	(0.0805)
	First Stage F-Stat		2.648
	Number of Obs	8,956	8,956
Number of New Products Per Store	Treat	1.403*	2.352*
		(0.828)	(1.354)
	Treat * Delivery	3.403	7.993
		(3.876)	(12.77)
	First Stage F-Stat		1.247
	Number of Obs	312	312
Store Owner Sources Products Online (Yes=1)	Treat	0.0250**	0.0416**
		(0.0122)	(0.0201)
	Treat * Delivery	-0.0911	-0.185
		(0.0814)	(0.166)
	First Stage F-Stat		1.320
	Number of Obs	341	341
Log Price of Business Inputs	Treat	-0.0858	-0.108
		(0.134)	(0.182)
	Treat * Delivery	0.289	0.473
		(0.273)	(0.447)
	First Stage F-Stat		1.972
	Number of Obs	237	237
Log Price of Non-Durables	Treat	0.0192	0.0366
		(0.0157)	(0.0308)
	Treat * Delivery	0.0137	0.0214
		(0.0362)	(0.0585)
	First Stage F-Stat		16.09
	Number of Obs	6,455	6,455
Log Prices of Durables	Treat	-0.118	-0.144
		(0.0880)	(0.104)
	Treat * Delivery	0.164	0.288
		(0.134)	(0.366)
	First Stage F-Stat		0.488
	Number of Obs	185	185

Notes: See Section 4 for discussion. Standard errors are clustered at the level of villages. * 10%, ** 5%, *** 1% significance levels.

Table 8: Heterogeneity Across Households and Villages

Type of Heterogeneity		Intent to Treat	Treatment on the Treated	Log Dist (IV)	Intent to Treat	Treatment on the Treated	Log Distance (IV)	Intent to Treat	Treatment on the Treated
	Dependent Variables:	Household Has Ever Bought Something at Terminal (Yes=1)			Monthly Income Per Capita (RMB)			Log Local Retail Prices	
Average Effect	Treat or Log Dist	0.0480***	0.0886***	-0.0253***	-7.838	-14.48	4.190	0.0189	0.0352
		(0.0169)	(0.0271)	(0.00801)	(70.78)	(129.9)	(37.55)	(0.0142)	(0.0263)
	R-Squared	0.008			0.038			0.893	0.893
	First Stage F-Stat		45.56	39.22		45.33	37.69		41.66
	Number of Obs	3,518	3,518	3,518	3,437	3,437	3,437	6,877	6,877
Village Was Previously Connected to Parcel Delivery (Yes=1)	Treat or Log Dist	0.0573***	0.105***	-0.0323***	-14.99	-27.14	8.513	0.0114	0.0215
		(0.0190)	(0.0288)	(0.00922)	(77.55)	(140.1)	(43.82)	(0.0144)	(0.0273)
	Treat or Log Dist *	-0.0603**	-0.110**	0.0335***	50.29	97.16	-22.44	0.0417	0.0739
	Delivery	(0.0251)	(0.0438)	(0.0113)	(171.2)	(339.1)	(75.42)	(0.0377)	(0.0572)
	R-Squared	0.016			0.040			0.894	
	First Stage F-Stat		2.683	14.88		2.694	14.42		17.26
	Number of Obs	3,518	3,518	3,518	3,437	3,437	3,437	6,877	6,877
Village Distance to Township Center	Treat or Log Dist	-0.0156	-0.00882	-0.00268	-23.53	-43.67	14.71	-0.0219	-0.0322
		(0.0288)	(0.0429)	(0.0126)	(181.7)	(289.2)	(84.33)	(0.0375)	(0.0632)
	Treat or Log Dist *	0.0388**	0.0612***	-0.0138**	0.389	0.371	-1.272	0.0216	0.0358
	Log Dist Planned	(0.0162)	(0.0227)	(0.00570)	(97.50)	(152.0)	(40.55)	(0.0198)	(0.0336)
	R-Squared	0.014			0.040			0.893	
	First Stage F-Stat		15.63	11.79		15.66	10.98		16.96
	Number of Obs	3,518	3,518	3,518	3,437	3,437	3,437	6,877	6,877
Primary Earner's Age	Treat or Log Dist	0.140***	0.223***	-0.0669***	-136.4	-237.8	70.34		
		(0.0506)	(0.0778)	(0.0230)	(172.5)	(286.5)	(84.03)		
	Treat or Log Dist *	-0.00172**	-0.00251*	0.000778**	2.561	4.551	-1.341		
	Age	(0.000774)	(0.00129)	(0.000370)	(2.734)	(4.825)	(1.404)		
	R-Squared	0.023			0.049				
	First Stage F-Stat		16.07	15.63		16.34	15.65		
	Number of Obs	3,304	3,304	3,304	3,292	3,292	3,292		
Primary Earner's Educaction	Treat or Log Dist	0.0407*	0.0977**	-0.0266**	52.80	119.7	-33.46		
		(0.0206)	(0.0412)	(0.0115)	(83.52)	(195.0)	(53.92)		
	Treat or Log Dist *	0.00161	-0.000469	-5.85e-05	-8.666	-17.79	5.057		
	Years of Education	(0.00267)	(0.00506)	(0.00141)	(12.14)	(24.03)	(6.774)		
	R-Squared	0.014			0.063				
	First Stage F-Stat		8.462	10.62		8.662	10.78		
	Number of Obs	3,296	3,296	3,296	3,284	3,284	3,284		
Household Income Per Capita	Treat or Log Dist	0.00806	0.0209	-0.00505	35.86	59.51	-16.75		
		(0.0213)	(0.0375)	(0.00998)	(96.83)	(165.5)	(45.62)		
	Treat or Log Dist *	0.00712**	0.0121**	-0.00370**	-9.204	-15.79	4.564		
	Log Income PC	(0.00326)	(0.00541)	(0.00162)	(21.22)	(36.31)	(10.39)		
	R-Squared	0.011			0.355				
	First Stage F-Stat		22.78	17.96		22.57	17.62		
	Number of Obs	3,416	3,416	3,416	3,437	3,437	3,437		
Household Distance to Planned Terminal	Treat or Log Dist	0.144**	0.231**	-0.0636**	185.9	400.1	-108.9		
		(0.0591)	(0.109)	(0.0315)	(350.6)	(697.5)	(188.3)		
	Treat or Log Dist *	-0.0181*	-0.0274	0.00739	-36.54	-79.67	21.85		
	Log Dist Planned	(0.00981)	(0.0193)	(0.00587)	(61.53)	(128.5)	(34.90)		
	R-Squared	0.012			0.039				
	First Stage F-Stat		9.905	11.64		9.325	14.15		
	Number of Obs	3,518	3,518	3,518	3,437	3,437	3,437		
Combined	Treat or Log Dist	0.154*	0.289**	-0.0838*	108.5	213.4	-57.26	-0.0398	-0.0435
		(0.0805)	(0.140)	(0.0438)	(333.8)	(619.5)	(184.7)	(0.0362)	(0.0531)
	Treat or Log Dist *	-0.0400	-0.106	0.0342**	98.21	229.2	-53.30	0.0413	0.0517
	Delivery	(0.0285)	(0.0687)	(0.0149)	(137.1)	(336.0)	(69.69)	(0.0361)	(0.0622)
	Treat or Log Dist *	0.0458***	0.0813***	-0.0178***	-37.85	-81.46	18.11	0.0284	0.0380
	Log Dist Township	(0.0174)	(0.0298)	(0.00688)	(62.90)	(134.2)	(31.65)	(0.0188)	(0.0312)
	Treat or Log Dist *	-0.00181**	-0.00314**	0.000964**	0.929	1.742	-0.511		
	Age	(0.000775)	(0.00129)	(0.000390)	(2.567)	(4.664)	(1.378)		
	Treat or Log Dist *	0.000370	-0.00380	0.000671	-2.778	-1.854	1.218		
	Years of Education	(0.00268)	(0.00499)	(0.00144)	(10.22)	(21.43)	(6.086)		
	Treat or Log Dist *	0.00908***	0.0162***	-0.00544***	-12.43	-21.38	6.717		
	Log Income PC	(0.00339)	(0.00555)	(0.00174)	(22.39)	(38.60)	(11.50)		
	Treat or Log Dist *	-0.0249**	-0.0417*	0.0109	-8.134	-20.40	5.556		
	Log Dist Planned	(0.0107)	(0.0218)	(0.00671)	(45.46)	(96.39)	(26.75)		
	R-Squared	0.051			0.353			0.894	
	First Stage F-Stat		0.474	2.991		0.420	2.938		1.579
	Number of Obs	3,261	3,261	3,261	3,282	3,282	3,282	6,877	6,877

Notes: See Section 4 for discussion. Standard errors are clustered at the level of villages. * 10%, ** 5%, *** 1% significance levels.

Table 9: How Do the E-Commerce Terminals Compare?

Could You Have Purchased This Product in Your Village? (Yes=1)	Sample Fraction	0.380	Household Living in Village Without Any Durables on Sale (Yes=1)	Sample Fraction	0.547
	Number Obs	255		Number Obs	3,508
Log Price Difference between Terminal and Village Retail	Sample Mean	-0.166	Travel Cost to Main Shopping Destination Outside Village (RMB)	Sample Mean	11.85
	Sample Median	-0.154		Sample Median	4
	Number Obs	95		Number Obs	2,766
Could You Have Purchased This Product in the Nearby Town? (Yes=1)	Sample Fraction	0.836	Travel Time to Main Shopping Destination Outside Village and Back (Minutes)	Sample Mean	58.14
				Sample Median	40
	Number Obs	238		Number Obs	2,366
Log Price Difference between Terminal and Nearby Town Retail	Sample Mean	-0.227	Travel Distance to Main Shopping Destination Outside Village and Back (Km)	Sample Mean	15.38
	Sample Median	-0.182		Sample Median	9.045
	Number Obs	197		Number Obs	2,773

Notes: See Section 4 for discussion.

Table 10: Role of GE Spillovers

Dependent Variables		Treatment on Treated without Spillovers	ToT with Spillovers: Number of Terminals within 3 km Outside of Village	ToT with Spillovers: Number of Terminals within 10 km Outside of Village
Monthly Income Per Capita (RMB)	Treat Dummy	-14.48	-3.924	-32.97
		(129.9)	(138.8)	(122.3)
	Exposure to Terminals Outside the Village		-143.9	-8.939
			(184.5)	(26.74)
	Exposure to Other Villages		-36.15**	-12.96***
			(15.91)	(3.917)
	First Stage F-Stat	45.33	47.82	44.55
	Number of Obs	3,437	3,437	3,437
Any Member of Household Has Ever Sold Online (Yes=1)	Treat Dummy	-0.0129	-0.0135	-0.0148
		(0.0104)	(0.0101)	(0.0101)
	Exposure to Terminals Outside the Village		-0.00142	-0.00233
			(0.0102)	(0.00202)
	Exposure to Other Villages		-0.00335***	-0.000285
			(0.00102)	(0.000363)
	First Stage F-Stat	45.30	47.63	44.61
	Number of Obs	3,504	3,504	3,504
Household Has Ever Bought Something at Terminal (Yes=1)	Treat Dummy	0.0886***	0.0786***	0.0862***
		(0.0271)	(0.0266)	(0.0267)
	Exposure to Terminals Outside the Village		0.0655**	-0.00611
			(0.0311)	(0.00568)
	Exposure to Other Villages		-0.00245	0.00252**
			(0.00538)	(0.00111)
	First Stage F-Stat	45.56	48.11	44.91
	Number of Obs	3,518	3,518	3,518
Share of Terminal in Total Retail Expenditure	Treat Dummy	0.0124***	0.0101**	0.0119***
		(0.00434)	(0.00398)	(0.00422)
	Exposure to Terminals Outside the Village		0.0159*	-0.00128
			(0.00834)	(0.000923)
	Exposure to Other Villages		-0.000594	0.000506**
			(0.000523)	(0.000228)
	First Stage F-Stat	44.03	46.57	43.50
	Number of Obs	3,434	3,434	3,434
Log Local Retail Prices (All Prices)	Treat Dummy	0.0352	0.0338	0.0386
		(0.0263)	(0.0258)	(0.0252)
	Exposure to Terminals Outside the Village		0.00353	0.00382
			(0.0314)	(0.00562)
	Exposure to Other Villages		-0.00318	-0.00135
			(0.00314)	(0.000950)
	First Stage F-Stat	41.66	43.89	43.95
	Number of Obs	6,877	6,877	6,877

Notes: See Section 4 for discussion. Standard errors are clustered at the level of villages. * 10%, ** 5%, *** 1% significance levels.

Table 11: Average Effects On Household Economic Welfare

	Unweighted (Effects in Sample)			Weighted (Effects in Village Population)		
	Durables Consumption	Non-Durables Consumption	Total Retail Consumption	Durables Consumption	Non-Durables Consumption	Total Retail Consumption
Reduction in Retail Cost of Living for All Households	3.298% (0.027)	0.478% (0.004)	0.812% (0.005)	2.908% (0.031)	0.419% (0.003)	0.714% (0.005)
Reduction in Retail Cost of Living Among Users	19.331% (0.215)	3.722% (0.029)	5.464% (0.035)	16.599% (0.215)	3.267% (0.024)	4.764% (0.032)

Notes: See Section 6 for discussion. Standard errors are bootstrapped across 1000 iterations with random re-sampling.

Appendix - For Online Publication

A Additional Figures and Tables

Table A.1: Extended Descriptive Statistics: Individual Level

		Full Sample at Baseline	Treatment Villages at Baseline	Control Villages at Baseline	P-Value (Treat-Control=0)	Control Villages at Endline
Age	Median	44.000	44.000	43.000		46.000
	Mean	38.950	39.329	38.407	0.208	39.943
	Standard Deviation	23.580	23.658	23.460		23.759
	Number of Obs	8491	5001	3490		4194
Gender (Female=1)	Median	1.000	1.000	1.000		1.000
	Mean	0.534	0.526	0.546	0.025	0.537
	Standard Deviation	0.499	0.499	0.498		0.499
	Number of Obs	8484	5001	3483		4188
Employed (for age>15) (Yes=1)	Median	1.000	1.000	1.000		1.000
	Mean	0.767	0.766	0.769	0.882	0.762
	Standard Deviation	0.423	0.424	0.422		0.426
	Number of Obs	6070	3590	2480		3015
Peasant (for age>15) (Yes=1)	Median	1.000	1.000	1.000		1.000
	Mean	0.527	0.527	0.526	0.971	0.513
	Standard Deviation	0.499	0.499	0.499		0.500
	Number of Obs	6369	3760	2609		3144
No Schooling (for age>15) (No School=1)	Median	0.000	0.000	0.000		0.000
	Mean	0.270	0.273	0.266	0.745	0.319
	Standard Deviation	0.444	0.446	0.442		0.466
	Number of Obs	6368	3758	2610		3132
Completed Junior High School (for age>15) (Yes=1)	Median	0.000	0.000	0.000		0.000
	Mean	0.437	0.429	0.449	0.419	0.422
	Standard Deviation	0.496	0.495	0.498		0.494
	Number of Obs	6368	3758	2610		3132
Completed Senior High School (for age>18) (Yes=1)	Median	0.000	0.000	0.000		0.000
	Mean	0.104	0.104	0.104	0.969	0.097
	Standard Deviation	0.305	0.305	0.305		0.296
	Number of Obs	6286	3719	2567		3096

Notes: See Section 2 for discussion.

Table A.2: Extended Descriptive Statistics: Household Level

		Full Sample at Baseline	Treatment Villages at Baseline	Control Villages at Baseline	P-Value (Treat-Control=0)	Control Villages at Endline
Age of Primary Earner	Median	50.000	50.000	50.000		52.00
	Mean	49.824	49.953	49.631	0.634	51.395
	Standard Deviation	12.673	12.710	12.621		13.55
	Number of Obs	2548	1530	1018		1348
Gender of Primary Earner (Female=1)	Median	0.000	0.000	0.000		0.00
	Mean	0.288	0.295	0.276	0.457	0.295
	Standard Deviation	0.453	0.456	0.447		0.46
	Number of Obs	2547	1530	1017		1348
Primary Earner Went to School (Yes=1)	Median	1.000	1.000	1.000		1.00
	Mean	0.815	0.814	0.817	0.874	0.750
	Standard Deviation	0.388	0.389	0.386		0.43
	Number of Obs	2550	1531	1019		1342
Primary Earner Is Peasant (Yes=1)	Median	1.000	1.000	1.000		1.00
	Mean	0.590	0.600	0.577	0.620	0.587
	Standard Deviation	0.492	0.490	0.494		0.49
	Number of Obs	2549	1531	1018		1348
Primary Earner Self-Employed (Yes=1)	Median	0.000	0.000	0.000		0.00
	Mean	0.073	0.087	0.053	0.036	0.072
	Standard Deviation	0.261	0.282	0.224		0.26
	Number of Obs	2549	1531	1018		1348
Household Size	Median	3.000	3.000	3.000		3.00
	Mean	3.114	3.053	3.205	0.075	2.987
	Standard Deviation	1.422	1.420	1.421		1.40
	Number of Obs	2740	1647	1093		1405
Household Monthly Income Per Capita in RMB	Median	350.000	339.000	375.000		466.67
	Mean	876.412	841.198	929.473	0.365	1028.960
	Standard Deviation	1717.456	1687.169	1761.560		2005.31
	Number of Obs	2740	1647	1093		1405
Household Monthly Retail Expenditure Per Capita in RMB	Median	381.000	372.833	400.500		364.00
	Mean	732.017	663.034	835.966	0.135	686.616
	Standard Deviation	2304.540	1139.788	3368.220		1512.06
	Number of Obs	2735	1644	1091		1405
Household Monthly Expenditure on Business Inputs Per Capita in RMB	Median	0.000	0.000	0.000		0.00
	Mean	123.417	123.007	124.033	0.981	128.464
	Standard Deviation	1033.757	1076.656	966.070		1069.52
	Number of Obs	2736	1644	1092		1405
Any Member of the Household Has Ever Used the Internet (Yes=1)	Median	0.000	0.000	0.000		0.00
	Mean	0.368	0.354	0.390	0.249	0.427
	Standard Deviation	0.482	0.478	0.488		0.49
	Number of Obs	2739	1646	1093		1402
Household Owns a Smartphone (Yes=1)	Median	1.000	1.000	1.000		1.00
	Mean	0.526	0.509	0.552	0.153	0.551
	Standard Deviation	0.499	0.500	0.498		0.50
	Number of Obs	2731	1642	1089		1400

Notes: See Section 2 for discussion.

Table A.3: Extended Descriptive Statistics: Household Level – Continued

		Full Sample at Baseline	Treatment Villages at Baseline	Control Villages at Baseline	P-Value (Treat-Control=0)	Control Villages at Endline
Share of Household Monthly Expenditure on E-Commerce Deliveries	Median	0.000	0.000	0.000		0.00
	Mean	0.007	0.006	0.007	0.693	0.008
	Standard Deviation	0.050	0.046	0.057		0.05
	Number of Obs	2720	1637	1083		1397
Share of E-Commerce Sales in Household Monthly Income	Median	0.000	0.000	0.000		0.00
	Mean	0.003	0.001	0.006	0.103	0.003
	Standard Deviation	0.052	0.030	0.074		0.05
	Number of Obs	2055	1244	811		1161
Distance in Meters to Planned Terminal Location	Median	231.556	232.891	231.454		203.63
	Mean	290.346	293.364	285.797	0.789	286.631
	Standard Deviation	243.450	247.778	236.820		267.06
	Number of Obs	2740	1647	1093		1405
Share of Retail Expenditure Outside of Village	Median	0.553	0.489	0.623		0.60
	Mean	0.500	0.470	0.545	0.193	0.531
	Standard Deviation	0.395	0.402	0.379		0.38
	Number of Obs	2720	1637	1083		1397
Share of Business Input Expenditure Outside of Village	Median	1.000	1.000	1.000		1.00
	Mean	0.613	0.610	0.618	0.916	0.633
	Standard Deviation	0.465	0.470	0.457		0.46
	Number of Obs	926	558	368		544
Travel Time One-Way to Main Shopping Destination Outside Village (minutes)	Median	20.000	20.000	20.000		20.00
	Mean	29.892	29.941	29.826	0.962	28.862
	Standard Deviation	27.825	27.380	28.429		26.19
	Number of Obs	2234	1284	950		1188
Travel Cost One-Way to Main Shopping Destination Outside Village (RMB)	Median	2.000	2.000	1.500		1.00
	Mean	3.739	3.847	3.591	0.715	4.236
	Standard Deviation	10.092	11.774	7.196		16.78
	Number of Obs	2216	1278	938		1185
Household Owns a PC or Laptop (Yes=1)	Median	0.000	0.000	0.000		0.00
	Mean	0.283	0.276	0.295	0.631	0.284
	Standard Deviation	0.451	0.447	0.456		0.45
	Number of Obs	2731	1642	1089		1400
Household Owns a Car (Yes=1)	Median	0.000	0.000	0.000		0.00
	Mean	0.108	0.107	0.110	0.851	0.131
	Standard Deviation	0.311	0.309	0.313		0.34
	Number of Obs	2731	1642	1089		1400
Household Owns a Motorcycle (Yes=1)	Median	0.000	0.000	1.000		0.00
	Mean	0.486	0.456	0.532	0.031	0.467
	Standard Deviation	0.500	0.498	0.499		0.50
	Number of Obs	2731	1642	1089		1400
Household Owns a TV (Yes=1)	Median	1.000	1.000	1.000		1.00
	Mean	0.977	0.977	0.977	0.953	0.977
	Standard Deviation	0.149	0.148	0.150		0.15
	Number of Obs	2731	1642	1089		1400

Notes: See Section 2 for discussion.

Table A.4: Extended Descriptive Statistics: Local Retail Prices

		Full Sample at Baseline	Treatment Villages at Baseline	Control Villages at Baseline	P-Value (Treat-Control=0)	Control Villages at Endline
Number of Stores at Village Level	Median	3.00	3.00	2.00		2.00
	Mean	4.15	4.38	3.79	0.33	3.61
	Standard Deviation	2.94	2.91	2.98		2.99
	Number of Obs	99	60	39		38
Establishment Space in Square Meters	Median	50.00	50.00	40.00		50.00
	Mean	99.07	74.42	146.76	0.35	121.33
	Standard Deviation	320.38	89.60	532.73		375.35
	Number of Obs	361	238	123		126
Number of Establishment's New Products Added Over Last Month	Median	0.00	0.00	0.00		0.00
	Mean	1.43	1.56	1.17	0.57	0.63
	Standard Deviation	7.44	8.88	3.42		2.26
	Number of Obs	330	215	115		126
Prices of All Retail Consumption (9 Product Groups) in RMB	Median	7.00	7.00	6.00		6.00
	Mean	71.03	76.74	61.43	0.47	71.23
	Standard Deviation	411.24	433.67	370.33		390.31
	Number of Obs	9382	5884	3498		3259
Price Was Not Displayed on Label (Needed to Ask=1)	Median	1.00	1.00	1.00		1.00
	Mean	0.67	0.66	0.67	0.97	0.73
	Standard Deviation	0.47	0.47	0.47		0.44
	Number of Obs	8977	5597	3380		3370
Prices of Business or Production Input in RMB	Median	10.00	10.00	8.80		9.00
	Mean	45.63	42.88	49.78	0.76	43.84
	Standard Deviation	195.09	206.23	177.46		97.92
	Number of Obs	444	267	177		111
(1) Prices of Food and Beverages in RMB	Median	4.38	4.60	4.00		4.00
	Mean	11.58	11.81	11.21	0.73	10.05
	Standard Deviation	24.35	23.31	25.99		17.75
	Number of Obs	4853	3021	1832		1834
(2) Prices of Tobacco and Alcohol in RMB	Median	12.00	13.00	12.00		13.00
	Mean	28.81	30.35	26.36	0.46	29.32
	Standard Deviation	53.97	59.45	43.77		55.16
	Number of Obs	1331	818	513		531
(3) Prices of Medicine and Health Products in RMB	Median	10.00	10.00	9.98		8.40
	Mean	26.13	24.40	29.31	0.66	18.50
	Standard Deviation	43.35	38.46	51.11		33.77
	Number of Obs	399	258	141		90
(4) Prices of Clothing and Accessories in RMB	Median	15.00	12.00	20.00		22.00
	Mean	46.31	45.69	47.79	0.90	57.00
	Standard Deviation	74.71	71.49	82.13		85.66
	Number of Obs	401	282	119		65
(5) Prices of Other Everyday Products in RMB	Median	10.00	10.00	9.00		9.00
	Mean	14.68	14.53	14.93	0.93	13.10
	Standard Deviation	31.03	32.69	28.06		18.17
	Number of Obs	1462	916	546		626
(6) Prices of Fuel and Gas in RMB	Median	5.00	5.00	5.00		5.83
	Mean	11.65	15.36	8.08	0.26	5.82
	Standard Deviation	21.46	28.88	9.59		0.23
	Number of Obs	53	26	27		4
(7) Prices of Furniture and Appliances in RMB	Median	110.00	85.00	187.00		398.00
	Mean	1009.49	1001.66	1026.34	0.95	1167.30
	Standard Deviation	1504.81	1583.03	1333.52		1350.70
	Number of Obs	183	125	58		43
(8) Prices of Electronics in RMB	Median	449.00	609.50	17.50		1799.00
	Mean	917.05	976.41	782.14	0.59	1782.71
	Standard Deviation	1224.37	1242.82	1184.20		871.58
	Number of Obs	144	100	44		45
(9) Prices of Transport Equipment in RMB	Median	1440.00	1980.00	30.00		2800.00
	Mean	1700.66	1794.74	1534.21	0.71	2578.24
	Standard Deviation	1822.07	1770.33	1922.34		1697.82
	Number of Obs	108	69	39		21

Notes: See Section 2 for discussion.

Table A.5: Test for Effects on Migration

Dependent Variables		Intent to Treat	Treatment on Treated	Log Distance (IV using Treat)
Attrition (Yes=1)	Treat or Log Dist	0.0138	0.0258	-0.00740
		(0.0239)	(0.0445)	(0.0127)
	R-Squared	0.000		
	Number of Obs	2,629	2,629	2,629
	First Stage F-Stat		44.24	35.90
Number of Household Members Who Moved Back to the Village	Treat or Log Dist	0.0255	0.0472	-0.0129
		(0.0400)	(0.0734)	(0.0199)
	R-Squared	0.001		
	Number of Obs	3,526	3,526	3,526
	First Stage F-Stat		45.27	42.71
Number of Household Members Who Moved Away from the Village	Treat or Log Dist	-0.00345	-0.00637	0.00174
		(0.0184)	(0.0338)	(0.00922)
	R-Squared	0.012		
	Number of Obs	3,523	3,523	3,523
	First Stage F-Stat		45.44	43.84
Would You Be Willing to Migrate to a City If a Good Job Opportunity Presented Itself? (Yes=1)	Treat or Log Dist	-0.0249	-0.0458	0.0125
		(0.0191)	(0.0348)	(0.00953)
	R-Squared	0.025		
	Number of Obs	3,527	3,527	3,527
	First Stage F-Stat		45.76	44.15

Notes: See Section 2 for discussion.

Table A.6: Role of Program Implementation

Type of Heterogeneity		Intent to Treat	Treatment on the Treated	Log Distance (IV Using Treat)
		Dependent Variable: Household Has Ever Bought Something at Terminal (Yes=1)		
Average Effects	Treat or Log Dist	0.0480***	0.0886***	-0.0241***
		(0.0169)	(0.0271)	(0.00721)
	R-Squared	0.008		
	First Stage F-Stat		45.56	43.80
	Number of Obs	3,518	3,518	3,518
Terminal Manager Test Score	Treat or Log Dist	0.0594	0.104	-0.0297
		(0.147)	(0.242)	(0.0679)
	Treat or Log Dist * Score	-0.000214	-0.000384	0.000114
		(0.00164)	(0.00270)	(0.000755)
	R-Squared	0.006		
	First Stage F-Stat		8.786	8.133
	Number of Obs	3,042	3,042	3,042
Terminal Manager Test Score Above the Median	Treat or Log Dist	0.0314	0.0616	-0.0172
		(0.0295)	(0.0501)	(0.0136)
	Treat or Log Dist * Above Median	0.0191	0.0182	-0.00504
		(0.0347)	(0.0583)	(0.0158)
	R-Squared	0.006		
	First Stage F-Stat		8.654	7.210
	Number of Obs	3,042	3,042	3,042
Terminal Installation Delayed	Treat or Log Dist	0.0392	0.0656*	-0.0180*
		(0.0247)	(0.0357)	(0.00941)
	Treat or Log Dist * Delay Dummy	0.0167	0.0486	-0.0131
		(0.0335)	(0.0554)	(0.0149)
	R-Squared	0.009		
	First Stage F-Stat		10.93	11.46
	Number of Obs	3,518	3,518	3,518

Notes: See Section 4 for discussion. Standard errors are clustered at the level of villages. * 10%, ** 5%, *** 1% significance levels.

Table A.7: Fraction of Market Access to Other Rural Markets in County

	Fraction of Market Access from Rural Markets in Same County						Fraction of Market Access from Participating Rural Markets in Same County					
Measure of Market Size:	Access to Population			Access to GDP			Access to Population			Access to GDP		
	Median	Mean	Std Dev	Median	Mean	Std Dev	Median	Mean	Std Dev	Median	Mean	Std Dev
Panel A: Distance Elasticity of -1												
All Rural Townships in East, Middle and Southwest China (10,214 Townships)	0.0082	0.011	0.01	0.0031	0.0044	0.005	0.0014	0.0018	0.0017	0.0005	0.0007	0.0008
Rural Townships in 3 RCT Provinces (2,291 Townships)	0.012	0.016	0.014	0.0037	0.0059	0.0062	0.0020	0.0027	0.0023	0.0006	0.0010	0.0010
Rural Townships in 8 RCT Counties (58 Townships)	0.011	0.012	0.006	0.0031	0.0041	0.0029	0.0018	0.0020	0.0010	0.0005	0.0007	0.0005
Panel A: Distance Elasticity of -1.5												
All Rural Townships in East, Middle and Southwest China (10,214 Townships)	0.027	0.037	0.042	0.01	0.016	0.024	0.0045	0.0062	0.0070	0.0017	0.0027	0.0040
Rural Townships in 3 RCT Provinces (2,291 Townships)	0.036	0.049	0.055	0.012	0.02	0.028	0.0060	0.0082	0.0092	0.0020	0.0033	0.0047
Rural Townships in 8 RCT Counties (58 Townships)	0.034	0.038	0.033	0.011	0.014	0.013	0.0057	0.0063	0.0055	0.0018	0.0023	0.0022

Notes: See Section 4.4 for discussion.

Table A.8: Are Sample Villages Representative?

	(1)	(2)	(3)	(4)	(5)	(6)
		Full Sample			3 Provinces	
Dependent Variables:	Number of Users	Number of Transactions	Sales (RMB)	Number of Users	Number of Transactions	Sales (RMB)
			Panel A: Purchase Database			
RCT_Sample Dummy	-4.110	0.0605	-6,034	0.149	12.65	-3,747
	(7.751)	(25.33)	(4,061)	(7.734)	(25.32)	(4,066)
Months Fixed Effects	✓	✓	✓	✓	✓	✓
Control for Months Since Program Entry	✓	✓	✓	✓	✓	✓
Observations	125,204	125,204	125,204	100,098	100,098	100,098
R-squared	0.037	0.047	0.029	0.031	0.046	0.03
Number of Village Clusters	11,731	11,731	11,731	8,471	8,471	8,471

	(7)	(8)	(9)	(10)
	Full Sample		3 Provinces	
Dependent Variables:	Number of Transactions	Weight (kg)	Number of Transactions	Weight (kg)
		Panel B: Out-Shipment Database		
RCT_Sample Dummy	1.712**	5.154	1.364*	4.68
	(0.753)	(4.332)	(0.752)	(4.333)
Months Fixed Effects	✓	✓	✓	✓
Control for Months Since Program Entry	✓	✓	✓	✓
Observations	120,483	120,483	95,744	95,744
R-squared	0.06	0.023	0.067	0.026
Number of Village Clusters	11,904	11,904	8,591	8,591

Notes: The upper panel presents point estimates from regressions based on the purchase transaction database. The lower panel presents point estimates from regressions based on the sales transaction database. See Section 5 for discussion. Standard errors are clustered at the level of village terminals. * 10%, ** 5%, *** 1% significance levels.

Table A.9: Role of Seasonality

	(1)	(2)	(3)	(4)	(5)	(6)
	Full Sample			3 Provinces		
Dependent Variables:	Number of Users	Number of Transactions	Sales (RMB)	Number of Users	Number of Transactions	Sales (RMB)
	Panel A: Purchase Database					
RCT Sample Month Dummy	0.893***	-4.671***	-1,565***	0.568**	-5.290***	-585.9
	(0.255)	(0.818)	(451.5)	(0.274)	(0.863)	(458.0)
Village Fixed Effects	✓	✓	✓	✓	✓	✓
Control for Months Since Program Entry	✓	✓	✓	✓	✓	✓
Observations	125,204	125,204	125,204	100,098	100,098	100,098
R-squared	0.694	0.68	0.219	0.679	0.667	0.227
Number of Village Clusters	11,731	11,731	11,731	8,471	8,471	8,471

	(7)	(8)	(9)	(10)
	Full Sample		3 Provinces	
Dependent Variables:	Number of Transactions	Weight (kg)	Number of Transactions	Weight (kg)
	Panel B: Out-Shipment Database			
RCT Sample Month Dummy	-0.387***	-1.256***	-0.498***	-1.407***
	(0.0225)	(0.125)	(0.0261)	(0.138)
Village Fixed Effects	✓	✓	✓	✓
Control for Months Since Program Entry	✓	✓	✓	✓
Observations	120,483	120,483	95,744	95,744
R-squared	0.592	0.432	0.57	0.422
Number of Village Clusters	11,904	11,904	8,591	8,591

Notes: The upper panel presents point estimates from regressions based on the purchase transaction database. The lower panel presents point estimates from regressions based on the sales transaction database. See Section 5 for discussion. Standard errors are clustered at the level of village terminals. * 10%, ** 5%, *** 1% significance levels.

Table A.10: Quantification Under Alternative Demand Parameters

	$\sigma_D = 2.87$, $\sigma_N = 2.85$			$\sigma_D = 3.87$, $\sigma_N = 3.85$			$\sigma_D = 4.87$, $\sigma_N = 4.85$		
	Durables Consumption	Non-Durables Consumption	Total Retail Consumption	Durables Consumption	Non-Durables Consumption	Total Retail Consumption	Durables Consumption	Non-Durables Consumption	Total Retail Consumption
Reduction in Retail Cost of Living for All Households	5.129% (0.043)	0.735% (0.005)	1.252% (0.007)	3.298% (0.027)	0.478% (0.004)	0.812% (0.005)	2.431% (0.02)	0.355% (0.003)	0.601% (0.003)
Reduction in Retail Cost of Living Among Users	31.47% (0.368)	5.773% (0.046)	8.526% (0.056)	19.331% (0.215)	3.722% (0.029)	5.464% (0.035)	13.942% (0.151)	2.747% (0.022)	4.02% (0.026)

Notes: See Section 6 for discussion. Standard errors are bootstrapped across 1000 iterations with random re-sampling.

B Theoretical Framework for Welfare Evaluation

Following recent work by (Atkin et al., in press), we propose a three-tier demand system to describe household retail consumption across product groups, retail shopping options and products. In the upper tier, shown in equation A.1, there are Cobb-Douglas preferences over broad product groups $g \in G$ (durables and non-durables) in total consumption. In the middle tier, shown in equation A.2, there are asymmetric CES preferences over local retailers selling that product group $s \in S$ (e.g. local stores, market stalls or the e-commerce terminal). In the final tier, there are preferences over the individual products within the product groups $b \in B_g$ that we can leave unspecified for now.

$$U_h = \prod_{g \in G} \left[Q_{gh}\right]^{\alpha_{gh}} \tag{A.1}$$

$$Q_{gh} = \Big(\sum_{s \in S_g} \beta_{gsh} q_{gsh}^{\frac{\sigma_g - 1}{\sigma_g}}\Big)^{\frac{\sigma_g}{\sigma_g - 1}} \tag{A.2}$$

where α_{gh} and β_{gsh} are (potentially household group-specific) preference parameters that are fixed across periods. Q_{gh} and q_{gsh} are product-group and store-product-group consumption aggregates with associated price indices P_{gh} and r_{gsh} respectively, and σ_g is the elasticity of substitution across local retail outlets. For each broad product group, consumers choose how much they are going to spend at different retail outlets based on the store-level price index r_{gsh} (which itself depends on the product mix and product-level prices on offer across outlets).

While the demand system is homothetic, we capture potential heterogeneity across the income distribution by allowing households of different incomes to differ in their expenditure shares across product groups (α_{gh}) and their preferences for consumption bundles at different stores within those product groups (β_{gsh} and the preference parameters that generate q_{gsh}). As shown by Anderson et al. (1992), these preferences can generate the same demands as would be obtained from aggregating many consumers who make discrete choices over which store to shop in. Building on Feenstra (1994), the following expression provides the exact proportional cost of living effect under this demand system:

$$\frac{CLE}{e(\mathbf{P}_T^{0*}, \mathbf{P}_C^{0}, \mathbf{P}_E^{0*}, \mathbf{P}_X^{0}, u_h^0)} = \frac{e(\mathbf{P}_T^{1}, \mathbf{P}_C^{1}, \mathbf{P}_E^{1}, \mathbf{P}_X^{1*}, u_h^0)}{e(\mathbf{P}_T^{0*}, \mathbf{P}_C^{0}, \mathbf{P}_E^{0*}, \mathbf{P}_X^{0}, u_h^0)} - 1 = \prod_{g \in G} \left(\Big(\frac{\sum_{s \in S_g^C} \phi_{gsh}^1}{\sum_{s \in S_g^C} \phi_{gsh}^0}\Big)^{\frac{1}{\sigma_g - 1}} \prod_{s \in S_g^C} \Big(\frac{r_{gsh}^1}{r_{gsh}^0}\Big)^{\omega_{gsh}} \right)^{\alpha_{gh}} \tag{A.3}$$

where S_g^C denotes the set of continuing local retailers within product group g, $\phi_{gsh}^t = r_{gsh}^t q_{gsh}^t / \sum_{s \in S_g} r_{gsh}^t q_{gsh}^t$ is the expenditure share for a particular retailer of product group g, and the ω_{gsh}s are ideal log-change weights.[1]

For each product group g, the expression has two components. The $\prod_{s \in S_g^C} (\frac{r_{gsh}^1}{r_{gsh}^0})^{\omega_{gsh}}$ term is a Sato-Vartia (i.e. CES) price-index for price changes in continuing local stores that forms the *pro-competitive price effect*.[2] The price terms r_{gsh}^t are themselves price indices of product-specific prices p_{gsb}^t within local continuing stores which, in principle, could also account for new product varieties or exiting product varieties using the same methodology. While we name these price changes pro-competitive, they may derive from either reductions in markups or increases in productivity at local stores (distinctions that do not matter on the cost-of-living side, but would

[1] In particular, $\omega_{gsh} = \left(\frac{\tilde{\phi}_{gsh}^1 - \tilde{\phi}_{gsh}^0}{\ln \tilde{\phi}_{gsh}^1 - \ln \tilde{\phi}_{gsh}^0}\right) / \sum_{s \in S_g^C} \left(\frac{\tilde{\phi}_{gsh}^1 - \tilde{\phi}_{gsh}^0}{\ln \tilde{\phi}_{gsh}^1 - \ln \tilde{\phi}_{gsh}^0}\right)$, which in turn contain expenditure shares of different retailers within product groups, where the shares consider only expenditure at continuing retailers $\tilde{\phi}_{gsh}^t = r_{gsh}^t q_{gsh}^t / \sum_{s \in S_g^C} r_{gsh}^t q_{gsh}^t$.

[2] Notice that the assumption of CES preferences does not imply the absence of pro-competitive effects as we do not impose additional assumptions about market structure (e.g. monopolistic competition).

generate different magnitudes of profit and income effects that we capture on the nominal income side).

The $\left(\frac{\sum_{s\in S_g^C}\phi^1_{gsh}}{\sum_{s\in S_g^C}\phi^0_{gsh}}\right)^{\frac{1}{\sigma_{gh}-1}}$ term captures the gains to customers of the e-commerce terminal in the numerator, from both the *direct price index effect* and the *entry effect*, and local store exit in the denominator, i.e. the *exit effect*. As in expression (2) of Section 3, we can decompose the total cost of living effect in equation (A.3) into four different types of effective consumer price changes by adding and subtracting terms.

Consider the case where the program's effect on cost of living is driven entirely by the direct price index effect. In that case, the expenditure share spent on continuing local retailers ($\sum_{s\in S_g^C}\phi^1_{gsh}$) is lower than unity only due substitution to the new e-commerce terminal (abstracting from a potential effect on store entry). The gains from the program as a proportion of initial household spending are then:

$$\frac{DE}{e(\mathbf{P}_T^{0*},\mathbf{P}_C^0,\mathbf{P}_E^{0*},\mathbf{P}_X^0,u_h^0)} = \prod_{g\in G}\left(\left(\sum_{s\in S_g^C}\phi^1_{gsh}\right)^{\frac{1}{\sigma_g-1}}\right)^{\alpha_{gh}} - 1. \tag{A.4}$$

The welfare gain from a new shopping option is a function of the market share of that outlet post-entry and the elasticity of substitution across stores. The revealed preference nature of this approach is clear. If consumers greatly value the arrival of the new option—be it because it offers low prices p^1_{gsb}, more product variety that reduces r^1_{gsh} or better amenities β_{gsh}—the market share is higher and the welfare gain greater. Hence, these market share changes capture all the potential consumer benefits of shopping through the e-commerce terminal. The magnitude of the welfare gain depends on the elasticity of substitution. Large terminal market shares will imply small welfare changes if consumers substitute between local shopping options very elastically, and large welfare changes if they are inelastic. A similar logic would apply to effects on the entry of local retailers, or on the exit of local stores (where a large period 0 market share means large welfare losses, again tempered by the elasticity of substitution).

C Data Appendix

C.1 Surveyor Training and Quality Management

This section describes our methodology for surveyor training and quality management. All survey and training material is available from the authors upon request.

Piloting and Surveyor Training Our survey supervisors are professionals from the Research Center for Contemporary China (RCCC) at Peking University. All RCCC supervisors have previous experience conducting large scale surveys in rural China. Before each of the two survey rounds, we traveled to Beijing to lead a one-day training workshop targeted at the supervisors and a group of graduate students from Renmin University and Jinan University, who were working with us as research assistants on this project. This training walked the RCCC supervisors and our graduate students through each step of the survey design, data collection protocols and quality control protocols that we had shared with them to study carefully in advance. Given budget and time constraints, the survey was paper based. Prior to our baseline survey, RCCC supervisors and our team of graduate students tested our survey design in a pilot survey of 45 households in two villages located in the rural parts of Beijing Province.

In the field, each supervisor was in charge of a team of six surveyors. In addition to the supervisors, two of our trained graduate students accompanied each team in the field. The role of the graduate students was to both support and monitor the recruitment and training of the local surveyors and the data collection, and to report back to us with detailed daily progress reports. Given differences in local dialects and rural conditions, the RCCC recruited surveyors among local university students from the provinces in which the data collection took place. All surveyors are familiar with the local dialect and customs of the rural areas in their home province. Each surveyor completed at least two full days of training and supervised practice questionnaire interviews before joining our field survey team. As part of the training, we provided surveyors with a number of supporting documents. In particular, they received an example of a completed representative survey questionnaire, detailed instructions on how to assist households in answering the questionnaire, a set of cards containing descriptions and examples of consumption products within categories or income-generating activities within sectors, and a set of solutions and best practices for common survey challenges. As described in Appendix C.4 below, we also trained surveyors to use separate pre-prepared spreadsheets to list individual household purchase transactions within product categories or income flows by type of activity. These spreadsheets were used for households to list individual transactions over a given period of time and within categories, before aggregating this information up to complete the final survey questionnaire cells. As part of their training, surveyors were trained to double-check with respondents any answer to the questionnaire that appears inconsistent with a previous answer.

Data Quality Management and Cleaning Surveyors conducted the household survey in pairs. During the interview, surveyors completed the questionnaire, along with supporting documents used to help households recall, categorize and sum up their consumption expenditures or earnings (we further describe data collection and variable construction for expenditure and earning variables in Appendix C.4 below). As part of quality control, supervisors reviewed one randomly chosen completed questionnaire, supporting documents, and interview audio tape from each surveyor at the end of every day.[3] In addition, our graduate students monitored the survey teams by accompanying them for part of their interviews, and reported back to the supervisors and our team in case of concerns. During recruiting and surveyor training, the surveyors had been informed that lack of accuracy, diligence or patience in the interviews would lead to the termination of employment, while a good record guaranteed a letter of recommendation confirming participation in our research project.

We also asked our surveyors to rate each household respondent along a number of dimensions such as cooperativeness, reliability, level of understanding, and level of interest in our survey. Surveyors also recorded the presence of any other household or non-household member whose presence could affect answers to our questionnaire. In our analysis of the data, we paid special attention to the reliability rating: 1. completely reliable, 2. mostly reliable, and 3. sometimes not reliable. Whenever surveyors rated a respondent as "sometimes not reliable", they also wrote down an explanation for this rating. On the basis of these written explanations, we created a clean household survey dataset. This dataset excludes 0.25 percent of unreliable/uncooperative households entirely from the sample. In other cases, surveyors' explanation suggested that only answers to a particular section of our questionnaire were unreliable. Using this information, we set all income variables to missing for 1.06 percent of all household respondents, all consumption variables to missing for 0.4 percent of households, and all income and consumption variables to missing for 1.31 percent of households. The descriptive statistics in Tables 1 and 2 and A.1-A.4 are based on this cleaned household survey dataset.

C.2 Household Sampling, Response Rates and Attrition

Our team was granted a two-week window for data collection, after receiving the extended candidate list of candidate villages from the local operation team in each county. Given this tight timeline, we were unable to conduct a village census for sampling purposes. Instead, our survey teams created detailed maps of all residences in the village to implement a random walk procedure. [4]

[3]Some households opted out of audio-recording.

[4]We use the boundary of the "natural village" as opposed to the "administrative village". Both of these are known delineations in Chinese. The natural village captures a geographically contiguous rural population. Administrative

From each village's map, we defined an "inner zone" of residences within a 300 meter radius of the planned terminal location and an "outer zone" outside that radius. In the baseline, the objective was to sample 14 households from the inner zone and 14 households from the outer zone. To randomly sample households within these zones, we selected 24 residences in both inner and outer zones. The household sampling proceeds as follows: we first randomly assign numbers to all residences within the zone on the map from 1 to n, and then define a rounded integer number $n/24$. Starting from household number 1, we then collect survey data from every household number in steps of the integer $n/24$ until we have completed 14 surveys within the zone. For the endline, we implement the same procedure for all households that were not part of the baseline survey to select 10 additional households within the inner zone. In the few case in which there were fewer than 24residences within the inner zone, we extended the radius until we obtain at least 24 residences on the map.

After introducing our survey to households, our surveyors asked for the household member with the best knowledge of household consumption expenses and household incomes to respond to the questionnaire. In case nobody answered the door, or in case this most suited household member was not at home during our surveyors' first visit, the surveyors returned at least twice to complete the interview, often outside of working hours, before moving on to the next household on the list. Surveyors were also instructed to skip households with a most knowledgeable respondent older than 75. Overall, our surveyors found willing and able respondents in two thirds of visited residences (66.1 percent).[5] In the endline, we sampled an additional 10 households from the inner zone. We used the same sampling methodology as in the baseline. Given expected sample attrition and the objective of 10 randomly selected additional households, the survey teams created a list of 22 new residential addresses in the inner zone and 6 new addresses in the outer zone. In the endline, we replaced a household respondent from the baseline whenever either the household had moved, the primary earner was no longer living there or the original baseline respondent was unavailable after three interview attempts. Using this rule, 71 percent of baseline respondents completed our questionnaire in the endline. As documented in Table A.5, this percentage does not differ in treatment and control villages.

C.3 Retail Price Survey

Store Sampling Prior to the field survey, RCCC supervisors performed a census of all retail stores and market stalls ("stores" for short) within a 15-minute walking distance of the boundaries of the natural village. Most villages have fewer than five stores, so in most villages we sampled products from all stores and market stalls in the vicinity of the village. If there were more than 15 stores in a village, we instructed supervisors to collect a representative sample of local retail information, giving more weight (i.e. more price quotes) to more popular establishments within product groups.

Product Sampling and Data Collection The data collection for the local retail price survey was conducted by the trained RCCC supervisors. We aim to collect data on 115 price quotes for each village. 100 of these prices are from the same 9 household consumption categories for retail products as in our household survey (food and beverages, tobacco and alcohol, medicine and health, clothing and accessories, other every-day products, fuel and gas, furniture and appliances, electronics, transport equipment), and 15 price quotes are for local production/business inputs. Our protocol for the price data collection closely follows the IMF/ILO standards for store price surveys that central banks collect to compute the CPI statistics. The sampling of products across consumption categories is based on budget shares of rural households in Anhui and Henan that

villages have a village committee. In most cases, the boundaries of the natural and administrative village correspond. In some cases, the administrative village includes more than one natural village.

[5]Of the one third of addresses that our surveyors did not encounter willing and able respondents, 56.6 percent had nobody at home during any of our three visits, 30.5 percent refused to participate in the survey, 7.5 percent had no qualified respondent (due to old age), and 6.4 percent had no one living there.

we observe in the microdata of the China Family Panel Study (CFPS) for 2012. Reflecting these consumption weights, supervisors in the baseline survey data aim to collect 47/100 price quotes in food and beverages, 15/100 in tobacco and alcohol, 9/100 in medicine and health, 9/100 in clothing and accessories, 4/100 in other every-day products, 4/100 in fuel and gas, 4/100 in furniture and appliances, 4/100 in electronics and 4/100 in transport equipment. In addition, we collect 15 price quotes for purchases of inputs to production or businesses.[6]

We provided supervisors with pre-structured price surveys reflecting the number of observations to be collected for each product group. As for the collection of data on household expenses that we discuss above and in Appendix C.4 below, the supervisors were provided with detailed product cards that list product groups within each of the 10 broad categories above, as well as examples of product types within those subgroups of products. They also received instructions on product sampling, for instance about how to evaluate the popularity of an individual product by measuring shelf space and recurrence across different stores. To ensure that we can match identical products in both survey rounds, supervisors saved a picture of each product and recorded product characteristics at the barcode-equivalent level, including packaging type, size, and a detailed product description (name, brand, flavor, etc) wherever possible.[7] For 78 percent of products collected in the baseline, we were able to find the exact same product in the same store one year later in the endline. As documented in Table 6, this percentage is somewhat smaller in intent to treat villages than in control villages, but this difference is not statistically significant. One challenge of surveying prices in rural China is a frequent lack of price tags displayed in store. As shown in Table A.4, about two thirds of the surveyed products lacked a price tag. In these cases, supervisors asked the store owner for the price that villagers would pay for the product. As part of quality control, we asked supervisors to rate the reliability of store owners' price quotes as good, average or poor. The reported findings in Tables 6-8 and 10 do not change in sign, size or statistical significance when limiting the sample to price quotes from labeled products only or excluding reportedly unreliable price quotes.

C.4 Variable Construction

To collect data on household consumption expenditures and incomes from different activities, we trained the surveyors in using separate pre-prepared spreadsheets before filling out the final survey questionnaires. For expenditures, there is one spreadsheet for each of the nine categories that we include in retail consumption, and a separate sheet for business inputs. This allowed households to recall and list all relevant expenses or income flows within a given product group or type of activity over a given period of time. This transaction-level information was then aggregated in the presence of the household to complete the final survey questionnaire sections on expenditures or income flows. To help respondent recall and categorize their expenditures, surveyors also received cards with examples of products in each category. The product cards break down the retail consumption space into 169 product types within the 10 broad categories we list above. After recording each item in a given category, surveyors go through the list of items and ask respondents how much they paid for each listed purchase. In addition to different consumption product groups, the surveyors also recorded the modality of each listed purchase transaction (e.g. online vs offline, in the village vs outside the village).

The above procedure was implemented covering a two-week time window for non-durable household consumption, and a three-month time window for durable goods categories. To obtain total monthly retail expenditure, we multiply the bi-weekly expenditure on non-durables by a factor of 2 and divide durable good expenditure by a factor of 3, and sum up across the

[6]Supervisors sometimes failed to find enough products in a given category within the village. This was often the case for the durable goods categories. In such cases, supervisors replaced products in these missing categories with additional price quotes for products in "other every-day used products".

[7]Some store owners refused to let supervisors take pictures. In such cases, we identify identical products in the endline data based on the same store and the detailed recorded product description.

9 consumption categories. For online expenditures at the terminal, we include both direct use of the terminal interface as well as remote usage by ordering deliveries to the terminal through the firm's app. The majority of terminal usage (more than 60 percent) are done in person at the terminal rather than remotely. In the vast majority of village cases, deliveries and pickups can be made at the terminal location (90 percent). In about 10 percent of cases, the logistics operators have offered delivery to the home address too.

To construct total household income, our surveyors again used a pre-prepared spreadsheet to assist households in recording each of their individual income sources over the last month. We defined four income categories: farm earnings, non-farm earnings, remittances (money or in-kind) from family not living in the home, and all other income (e.g., pension, returns from savings, gifts). In addition, we recorded sector of activity and occupation categories for each economically active member of the household. To help household respondents recall and categorize earnings, surveyors used cards with detailed examples of income sources in each category and proceeded to collect each flow on the spreadsheet before filling out the final survey questionnaire in the presence of the household. Our measure of income per capita is the sum of all income sources in these four categories, divided by the number of household members. Our measure of income net of transfers subtracts gifts and remittances from family not living in the home. Our measure of income per capita net of costs subtracts the recorded household expenses used to generate the reported flows of income.[8] The income variables exclude the market value of home production for own consumption.[9] Including this as part of household income has no effect on the statistical zeros that we report in the analysis.

When using total nominal retail expenditure or income in RMB as the dependent variable on the left-hand side of the regression, we censor these reported values at the one-percent level from the left and right tails within the survey round.[10] The point estimates remain statistical zeros in all cases, as is the case post-censoring, but the standard errors slightly increase.

[8]Remittances represent on average 13 percent of total household income in our sample. The distribution of this share is skewed, with most households reporting no remittances in our survey month, and a few households receiving a large share of their income through remittances.

[9]The market value of all food and beverages that the household produces for its own consumption amounts to on average less than 10 percent of household incomes.

[10]Given more than one percent of observations report zero incomes, nominal incomes are only censored at the 99-percent level from the right tail.