AER: Insights 2021, 3(1): 35–50
https://doi.org/10.1257/aeri.20190382

Connecting the Countryside via E-Commerce: Evidence from China†

By Victor Couture, Benjamin Faber, Yizhen Gu, and Lizhi Liu*

This paper estimates the impact of the first nationwide e-commerce expansion program on rural households. To do so, we combine a randomized control trial with new survey and administrative microdata. In contrast to existing case studies, we find little evidence for income gains to rural producers and workers. Instead, the gains are driven by a reduction in cost of living for a minority of rural households that tend to be younger, richer, and in more remote markets. These effects are mainly due to overcoming logistical barriers to e-commerce rather than additional investments to adapt e-commerce to the rural population. (*JEL* I31, L81, O12, O18, P25, P36)

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 United States as the largest e-commerce market.[1] Most of this growth has taken place in cities, but the Chinese government recently announced the expansion of e-commerce to the countryside as a national policy priority. The objective is to foster rural economic development and reduce the rural-urban economic divide.[2] Other developing countries with large rural populations, such as Egypt, India, and Vietnam, have recently announced similar e-commerce expansion plans.[3]

These policies have been motivated by a growing number of case studies on highly successful "e-commerce villages" that have experienced rapid output growth

* Couture: Vancouver School of Economics, University of British Columbia (email: victor.couture@ubc.ca); Faber: Department of Economics, UC Berkeley, and NBER (email: benfaber@econ.berkeley.edu); Gu: HSBC Business School, Peking University (email: yizhengu@phbs.pku.edu.cn); Liu: McDonough School of Business, Georgetown University (email: lizhi.liu@georgetown.edu). Rohini Pande was coeditor for this article. We are grateful to the Bill and Melinda Gates Foundation, CEGA, the Clausen Center, the Haas School of Business, the Weiss Family Fund, and the Natural Science Foundation of Guangdong Province, China (Grant No. 2020A1515011163) for providing funding for this study. We have benefited from outstanding research assistance by Hero Ashman, Allison Green, 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 Hongbing Gao, Zhenzhong Sheng, Liang Chen, Wentao Zhang, and Zhengwei Jiang for providing us access to and assistance with the transaction database. This research was approved by the Berkeley Office for the Protection of Human Subjects under IRB Protocol No. 2015-09-7944, and the research design was pre-registered under RCT ID AEARCTR-0001582. All views expressed and errors are our own.

† Go to https://doi.org/10.1257/aeri.20190382 to visit the article page for additional materials and author disclosure statement(s).

[1] This is in both number of users and total sales. See, e.g., PFSweb (2016) and Statista (2016).

[2] Alleviating poverty through rural e-commerce has featured in the government's *No.1 Central Document* each year since 2014.

[3] See, e.g., Egypt Ministry of Communications and Information Technology 2016, India Ministry of Electronics and Information Technology 2016, Prime Minister of Vietnam 2016, and UNCTAD's new technical assistance platform, "eTrade for All: Unlocking the Potential of E-Commerce in Developing Countries" (UNCTAD 2016).

by selling both agricultural and nonagricultural products to urban markets via e-commerce. One of the most prominent examples is China: by 2018, the largest e-commerce platform, Taobao, had branded more than 3,000 rural marketplaces as "Taobao villages" based on their high concentration of online sales (AliResearch 2018).[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, e-commerce is meant to raise rural incomes through higher demand for local production, better access to inputs, and stronger incentives for rural entrepreneurship. There has been 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 large consumer gains in rural areas.[5]

The recent growth of e-commerce in a number of rural markets has captured the imagination of policymakers, but important questions remain about whether market integration through online trading platforms can have a broad and significant impact on rural development. There is also little evidence on the characteristics of households and markets that may benefit more or less from e-commerce and on the effectiveness of investments targeted at lifting different types of barriers to rural e-commerce access.[6] To answer these questions, this paper studies the first nationwide e-commerce expansion program. From 2014 to 2018, this program connected more than 40,000 Chinese villages to e-commerce. Our analysis combines a randomized control trial (RCT), which we implemented 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.

E-commerce is the ability to buy and sell products through online transactions coupled with transport logistics for local parcel delivery and pickup from producers. Bringing e-commerce to the countryside in developing countries requires more than internet access. The internet is already available in most of the Chinese countryside due to both smartphones and expanding broadband access. Instead, there are two current barriers to rural e-commerce trading, which we refer to as the logistical and the transactional barriers. The logistical barrier relates to the lack of modern commercial parcel delivery services. These providers already operate distribution networks across Chinese cities but have not entered large parts of the countryside. One well-known challenge to rural transport logistics is the so-called "last mile" between urban logistical hubs and small pockets of rural population. The transactional barrier refers to the potential lack of familiarity with navigating online platforms or access to online payment methods that rural households may face. Villagers

[4] See, e.g., World Bank publications by Luo and Niu (2019) and Luo, Wang, and Zhang (2019). E-commerce villages have also received widespread media attention (e.g., "How Is Internet Shopping Changing Rural Villages in China?: Online Shopping in Rural China," *BBC*, 2015, https://www.bbc.co.uk/programmes/p033z2yw/p033yssy; Connor 2016; Freedman 2017; and Weller 2017).

[5] In the United States, the share of e-commerce in 2015 retail sales was about 10–15 percent (Federal Reserve Bank of St. Louis 2016). In China, Dobbs et al. (2013) report this share to be as high as 20–30 percent in smaller cities, and Fan et al. (2018) find this share increases by 1.2 percentage points as city population decreases by 10 percent.

[6] These questions complement the recent literature on the consumer gains from e-commerce in the United States (e.g., Brynjolfsson, Hu, and Smith 2003; Goldmanis et al. 2010; Dolfen et al. 2017).

may also 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/pickup near the urban center and fully subsidizes transport between the county's city center to and from the participating villages. To address additional transactional barriers specific to the rural population, the program installs an e-commerce terminal in a central village location. A terminal manager employed by the firm is available to assist villagers in buying and selling products through the firm's e-commerce platform. Villagers can pay upon receipt of their products or get paid upon pickup of their shipments in cash at the terminal location. The terminal is available in addition to the platform's online app-based interface for buying and selling.

An advantage of this setting is that we can study the reduction in trading frictions through e-commerce without confounding the counterfactual with the effects of first-time internet access or reductions in transport costs 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, for example, to control villages, remain unchanged.[7] 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.[8] In addition to evaluating the program's overall impact, we use the features of this setting to provide evidence on the relative importance of trade cost reductions (logistical barrier) and additional investments targeted at adapting e-commerce to the rural population (transactional barrier).

The analysis proceeds in two steps. In the first step, we randomize the arrival of e-commerce across 100 villages in 3 provinces and 8 counties and use our survey microdata to estimate the impact on local economic outcomes. We then bring to bear the firm's internal database covering the universe of transactions for about 12,000 villages in 5 provinces where the program had entered by April 2017. These data allow us to provide additional evidence on a number of questions outside the scope of the fieldwork. In particular, we investigate whether consumption or production-side effects take longer to materialize than the 12-month window we are able to study in the experiment and whether our household survey data may have missed rare but highly successful tail events on the producer side.

We interpret these results through the lens of a simple theoretical framework to quantify their implications for household welfare. We find no evidence of significant

[7] In this way we relate to but also differ from existing literature on the effects of transport cost reductions on rural markets (e.g., van de Walle 2009; Casaburi, Glennerster, and Suri 2013; Asher and Novosad 2020) and on the effects of the internet on rural markets (e.g., Chapman and Slaymaker 2002; Goyal 2010; Forman, Goldfarb, and Greenstein 2012; World Bank 2016). The empirical context and RCT allow us to study a different counterfactual of recent policy interest.

[8] We do not also attempt a social cost-benefit analysis of this program, which would require additional detailed and confidential information on the cost side from both the e-commerce firm and local and national governments, to which we do not have access.

gains or losses on the production and income sides of the local economy. This finding remains when using the firm's database to quantify village out-shipments up to 2.5 years after program arrival and using the universe of transaction records instead of survey samples. Instead, we find that the gains from e-commerce are driven by a reduction in cost of living for retail consumption. This effect is sizable (5 percent) among the group of rural households that are induced to use the new e-commerce option. These users, however, only represent about 15 percent of rural households, which are on average richer, younger, and living in more remote markets. In terms of channels, we find that the gains are concentrated among villages that were not previously serviced by commercial parcel delivery, suggesting the program's effects are mainly due to overcoming the logistical barrier rather than additional investments to lift transactional barriers specific to rural households. Consumer gains are strongest for durable product groups, such as electronics and appliances. We also find suggestive evidence of additional product variety in local stores, from sourcing new products through e-commerce. However, we find no evidence of procompetitive effects on local store prices for preexisting merchandise.

Overall, our findings put into context the transformative effects of e-commerce on rural markets that have been documented in numerous case studies on e-commerce villages in China and elsewhere. Our results suggest that such success stories are not representative of the countryside as a whole and should not be used as a guide to set policy expectations. Adding to this insight the significant heterogeneity that we document on the consumption side, access to e-commerce appears to offer economic gains to certain groups of the rural population and in certain places rather than being broad-based. As this evidence is based on one of the first and so far largest e-commerce expansion policies in the developing world, these findings are particularly relevant for the growing number of governments that have recently announced similar plans using China as a blueprint.[9] In this light, we hope that our work inspires future research aimed at investigating the local factors and potential complementary interventions, such as, for example, business training for e-commerce or access to credit, that enable certain groups and places to reap the gains from trade through e-commerce.

I. Experimental Design and Data

The experiment takes place in eight counties located in Anhui, Henan, and Guizhou provinces. The unit of randomization is the village. For each county, we obtain a list of villages where the firm plans to introduce the e-commerce program. We ask the firm to extend this list by five suitable village candidates in the county that would not have been part of the list in the absence of our research. We then randomly select five control villages and seven to eight treatment villages per county from this extended list. The remaining villages receive the e-commerce program as planned. The full sample in which we collect survey data thus includes 40 control villages and 60 treatment villages, randomly selected from 432 village candidates. Compliance with our assignments is not complete: the program was rolled out in 38

[9] In addition to the country plans discussed above, Thailand's recent "smart village" program has been designed based on field visits to Taobao villages in China ("eCommerce Ministry Touts Taobao Model," *Bangkok Post*, December 24, 2018, https://www.bangkokpost.com/business/1599750/commerce-ministry-touts-taobao-model).

of the 60 treatment villages and in 5 of the 40 control villages. We therefore report both intent-to-treat and treatment-on-treated effects. The main reason for imperfect compliance is that we are able to randomize treatments before the terminal manager applicants receive job offers, and some candidates end up rejecting.[10] Finally, in one of the counties, the local government suspended our team's data collection midway, leaving 4 of the 100 villages without endline data. The online Appendix provides additional details, maps, and descriptive statistics discussed below.

Household Survey Data.—For the baseline survey at the end of 2015 and beginning of 2016, we collect data from 28 households per village. Fourteen 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 second round of data collection occurs one year after the baseline.[11] We collect data from the same households as in the first round and were also able to extend the original sample by ten randomly sampled households within the inner zone. We collect detailed information about household retail consumption expenditures split across nine categories and for production and business inputs. We also collect information on household incomes, hours worked, occupations and sectors of different members, asset ownership, financial accounts, internet use, and migration.

The median age of all household members in the baseline survey is 44, and the median household size is 3. The primary earner is a farmer in 60 percent of households, and 82 percent of them completed at least primary school. Rural households are significantly poorer than in urban China: mean monthly income and retail expenditure per capita are about ¥876 and ¥732 respectively. Eighty percent of primary earners work inside the village. However, on average half of household retail expenditures occur outside the village, requiring a round trip to the nearest township center that takes on average one hour. Close to 40 percent of households report having used the internet, more than 50 percent own smartphones, and close to 30 percent report owning a laptop or personal computer. Almost all households own a television. At the same time, e-commerce penetration is very limited compared to urban regions: both the average share of household retail expenditure on e-commerce deliveries and the share of revenues from online selling in monthly income are less than 1 percent. Neither of these change over time in the endline survey among control villages.

Local Retail Price Survey Data.—We aim to collect 115 price quotes in each village. We sample products across nine retail consumption categories based on expenditure shares of rural households in Anhui and Henan from the 2012 China Family Panel Study (CFPS). We also include production and business inputs. We sample stores to be representative of local retail outlets (stores and market stalls). In villages with few stores, we sample all of them. We sample products within stores to capture a representative selection of locally purchased items within that store and

[10] Incomplete acceptance rates are standard in this setting and unrelated to the experiment (as applicants were unaware).

[11] The fast pace of the program's expansion places bounds on the timing of the endline. Our control villages ranked highly when the firm decided to launch additional waves of program expansion that were rolled out shortly after the endline.

product group. Each price quote is at the barcode-equivalent level when possible (recording brand, product name, packaging type, size, flavor if applicable). In the endline survey, we collect price quotes of the same products and retail outlets. In cases of either store closure or product disappearance, we include a new price quote within the same product category. The median number of sampled stores is three per village. The median floor space is 50 square meters, and the median store has not added new products within the last month.

Firm's Administrative Database.—We complement the survey data with administrative records from two different divisions of the firm covering five provinces (the three RCT provinces plus Guangxi and Yunnan, where the firm was also active). The first database covers the universe of e-commerce purchases made through the program in every participating village from November 2015 to April 2017. The data cover approximately 27.3 million transaction records across 12,000 villages over this period. For each transaction, the database contains information about the product category, number of units, amount paid, and a unique buyer identifier. Given that many villages had already been in operation for several months prior to November 2015, these data cover adjustment periods beyond the 12-months window that our RCT captures: transactions are observed up to 2 years and 4 months post-installation. The second database covers the universe of sales transactions—that is, out-shipments from the villages—through the firm's distribution network for the same universe of roughly 12,000 villages in the 5 provinces from January 2016 to April 2017. For each transaction, the database records the village of origin and the weight of the out-shipment in kilograms. The total number of e-commerce out-shipments over this period is roughly 500,000.

II. Analysis

A. *Evidence from Survey Data*

We run regressions of the following form:

$$y_{hv}^{Post} = \alpha + \beta_1 Treat_v + \gamma y_{hv}^{Pre} + \epsilon_{hv}, \tag{1}$$

where y_{hv} is an outcome of interest for household h living in village v.[12] For outcomes from the retail price data, h indexes individual price quotes or store-level outcomes instead. The variable $Treat_v$ is either an indicator of randomly assigned treatment status when estimating the intent-to-treat effect (ITT) or actual treatment status when estimating the treatment-on-the-treated effect (TOT) and instrumenting with intended treatment. We cluster standard errors at the level of the treatment (village level) and report point estimates both individually and after combining outcomes into category indices following Kling, Liebman, and Katz (2007) (KLK).

Table 1 presents estimation results for the average effects on household consumption (panel A), incomes (panel B) and local retail prices (panel C). Our

[12] While improving precision, none of the significant findings below rely on the inclusion of baseline outcomes y_{hv}^{Pre}.

Table 1—Average Effects

	ITT	TOT	ITT	TOT	ITT	TOT	ITT	TOT
Panel A. Consumption								
	Monthly retail expenditure per capita in renminbi		Has bought something through e-commerce option (yes = 1)		Share of e-commerce option in monthly total retail expenditure		Share of e-commerce option in monthly durables expenditure	
Treat	−22.09	−41.20	0.0484	0.0894	0.00668	0.0124	0.0408	0.0686
	(31.99)	(60.22)	(0.0167)	(0.0268)	(0.00239)	(0.00435)	(0.0160)	(0.0263)
R^2	0.038		0.008		0.006		0.012	
Control mean	592.21		0.0501		0.00277		0.0152	
First-stage F-statistic		44.01		45.31		44.03		52.43
Observations	3,436	3,436	3,518	3,518	3,434	3,434	768	768
Panel A. Consumption (continued)					*Panel B. Nominal incomes*			
	Share of e-commerce option in monthly nondurables expenditure		Consumption effects (KLK index)		Monthly income per capita in renminbi		Income effects (KLK index)	
Treat	0.00538	0.01	0.478	0.885	−7.864	−14.53	−0.0309	−0.0572
	(0.00196)	(0.00356)	(0.0336)	(0.126)	(70.78)	(129.9)	(0.0349)	(0.0646)
R^2	0.003		0.118		0.038		0.002	
Control mean	0.0027		0.00		915.51		0.00	
First-stage F-statistic		44.11		44.94		45.33		45.01
Observations	3,433	3,433	3,539	3,539	3,437	3,437	3,538	3,538
Panel C. Local retail prices								
	log prices		Product replacement dummy		Product addition dummy		Price effects (KLK index)	
Treat	0.0189	0.0352	−0.00392	−0.00747	2.194	4.020	−0.217	−0.389
	(0.0142)	(0.0263)	(0.0300)	(0.0569)	(1.073)	(2.278)	(0.134)	(0.260)
R^2	0.893		0.00		0.277		0.010	
Control mean	1.9813		0.0828		0.626		0.00	
First-stage F-statistic		41.66		39.82		19.69		24.05
Observations	6,877	6,877	8,956	8,956	312	312	343	343

Notes: Table reports point estimates from specification (1). Outcomes in panels A and B are at the household level. KLK consumption index based on 11 variables related to substitution into e-commerce, all entering positively (reducing price index). KLK income index based on 14 variables related to income generation, 13 entering positively and one negatively. In panel C, the first four columns are at the individual product item level. The final four columns are at the store level. KLK retail index based on four store-level variables, with two entering positively (reducing price index) and two negatively. See Section IIA for discussion. Standard errors are clustered at the level of villages.

discussion here focuses on the TOT results. On the consumption side, we find that the program leads to an uptake of on average 9 percent of households using the new e-commerce option in treatment villages compared to control villages. As documented by the nonzero mean among control villages, this effect masks additional uptake due to users in nearby control villages, increasing the effect on uptake to about 14 percent of village households. We further investigate such spillovers at the end of this section. The treatment effect on the new option's share in total household retail expenditure is 1.24 percent for the average village household. Thus, households that report having used the e-commerce option spent on average $0.0124/0.089 = 14.1$ percent of their retail consumption during the past month. We find stronger effects on durables compared to nondurables. For durables, the share of household expenditure is 6.9 percent for the average household, indicating a 45 percent shift in durable consumption to the new e-commerce option among uptaking households.[13] For nondurables, the treatment effect on the share of household retail expenditure is 1 percent for the average household, indicating that ever-users spend on average about 11 percent of total nondurables expenditure on the new e-commerce option. While households do shift part of their expenditures to e-commerce, there are no significant treatment effects on total monthly retail expenditures. The last column of Table 1, panel A, combines 11 outcomes related to substitution into e-commerce into a single index, defined as the equally weighted average of z-scores that are calculated by subtracting the mean and dividing by the standard deviation of the control group. The treatment effect on this index is 0.89 and significant at the 1 percent level.[14]

Table 1, panel B, reports point estimates on incomes per capita that are close to zero and not statistically significant. As above, we also report a single income-related index combining 14 outcomes related to income generation. We find no effects on either annual or monthly incomes, from agricultural or nonagricultural sources, on labor supply as measured by hours worked by the primary (or secondary) earner or on online selling activity, online revenues, sourcing of business inputs, or business creation offline or online. In terms of precision, the ITT point estimate on the income index indicates detectable positive effects down to about 2.6 percent of a standard deviation (one-sided 95 percent CI).

In Table 1, panel C, we find no significant reduction in local store prices for 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, the unweighted average effect on local retail prices is akin a Laspeyres price index for local retail consumption. We also find no effect when combining four outcomes related to local retail prices and product exit/additions into a single index. We find one piece of evidence suggestive of knock-on effects on preexisting local stores. The effect on the number of new products per store over the past month is four goods and is significant at the 10 percent level.

Heterogeneity.—In Table 2 we explore the heterogeneity of these effects. We begin by investigating the effect of the program as a function of preexisting availability of commercial parcel delivery at the village level. Villages serviced by commercial parcel delivery operators during our baseline survey already had access to local e-commerce deliveries. Interacting the treatment with preexisting parcel delivery status therefore allows us to shed light on the combined effect of removing both logistical and transactional barriers (among villages without preexisting parcel delivery) from the effect of removing only the transactional barrier (adding a terminal interface in villages with preexisting parcel delivery).[15] Next, we investigate heterogeneity across a basic set of household demographics that have been documented in

[13] For households that purchased durables over the past three months, the treatment effect on uptake is 15.3 percent instead of 9 percent. This yields an effect on the average durables consumption share among uptakers of $0.069/0.153 = 45$ percent.

[14] See online Appendix B for details on the KLK indices in Table 1.

[15] The transport subsidy does not affect villages previously serviced by parcel delivery, as logistics operators offered service in a few rural locations at the same rate as elsewhere in the county prior to program entry.

Table 2—Heterogeneity across Households and Villages

Type of heterogeneity	Household has bought something through e-commerce option (yes = 1) ITT	TOT	Monthly income per capita in renminbi ITT	TOT	log local retail prices ITT	TOT
Panel A. Village was previously connected to parcel delivery (yes = 1)						
Treat	0.0578	0.106	−15.00	−27.15	0.0114	0.0215
	(0.0188)	(0.0283)	(77.55)	(140.1)	(0.0144)	(0.0273)
Treat × delivery	−0.0606	−0.111	50.17	96.91	0.0417	0.0739
	(0.0253)	(0.0443)	(171.1)	(339.0)	(0.0377)	(0.0572)
First-stage F-statistic		2.682		2.694		17.26
Panel B. Village distance to township center						
Treat	−0.0144	−0.00652	−23.61	−43.80	−0.0219	−0.0322
	(0.0281)	(0.0411)	(181.7)	(289.1)	(0.0375)	(0.0632)
Treat × log dist. township	0.0384	0.0606	0.422	0.422	0.0216	0.0358
	(0.0161)	(0.0223)	(97.49)	(152.0)	(0.0198)	(0.0336)
First-stage F-statistic		15.55		15.66		16.96
Panel C. Primary earner's age						
Treat	0.141	0.223	−136.5	−238.0		
	(0.0505)	(0.0777)	(172.5)	(286.5)		
Treat × age	−0.00172	−0.00251	2.563	4.554		
	(0.000773)	(0.00129)	(2.734)	(4.825)		
First-stage F-statistic		16.04		16.34		
Panel D. Primary earner's education						
Treat	0.0408	0.0979	52.81	119.7		
	(0.0206)	(0.0412)	(83.52)	(195.0)		
Treat × years of education	0.00164	−0.000432	−8.672	−17.80		
	(0.00266)	(0.00504)	(12.14)	(24.03)		
First-stage F-statistic		8.456		8.662		
Panel E. Household income per capita						
Treat	0.00863	0.0220	35.83	59.45		
	(0.0214)	(0.0375)	(96.84)	(165.5)		
Treat × log income	0.00708	0.0120	−9.201	−15.78		
	(0.00327)	(0.00544)	(21.22)	(36.32)		
First-stage F-statistic		22.67		22.57		
Panel F. Household distance to planned terminal						
Treat	0.142	0.227	185.8	400.0		
	(0.0600)	(0.110)	(350.6)	(697.5)		
Treat × log dist. terminal	−0.0177	−0.0264	−36.53	−79.65		
	(0.0100)	(0.0196)	(61.53)	(128.5)		
First-stage F-statistic		9.899		9.325		
Panel G. Combined						
Treat	0.153	0.287	174.5	330.1	−0.0398	−0.0435
	(0.0811)	(0.141)	(329.9)	(612.1)	(0.0362)	(0.0531)
Treat × delivery	−0.0401	−0.106	102.1	253.3	0.0413	0.0517
	(0.0286)	(0.0690)	(121.1)	(308.1)	(0.0361)	(0.0622)
Treat × log dist. township	0.0457	0.0809	−42.86	−93.17	0.0284	0.0380
	(0.0173)	(0.0296)	(58.39)	(128.5)	(0.0188)	(0.0312)
Treat × age	−0.00181	−0.00314	0.587	1.266		
	(0.000775)	(0.00130)	(2.555)	(4.602)		
Treat × years of education	0.000384	−0.00377	−2.230	−1.954		
	(0.00267)	(0.00497)	(10.01)	(21.43)		
Treat × log income	0.00907	0.0162	−8.451	−14.28		
	(0.00339)	(0.00556)	(22.00)	(37.97)		
Treat × log dist. terminal	−0.0248	−0.0411	−16.48	−34.37		
	(0.0109)	(0.0222)	(45.01)	(94.93)		
First-stage F-statistic		0.479		0.419		1.579

Notes: Based on the same samples as Table 1. See Section IIA for discussion. Standard errors are clustered at the village level.

recent studies of internet and e-commerce use in China (respondent age, education, and income per capita) (China Internet Network Information Center 2015a, b). We also consider residential distance to the planned terminal location and a measure of village remoteness (motivated by Fan et al. 2018) based on road travel distance to the nearest township center. One should note that these interaction terms are not causally identified by experimental variation and provide additional suggestive evidence.

We first run regressions in which one characteristic at a time is interacted with the treatment, then a combined regression with all interactions included jointly. On the consumption side, we find that the effect on program uptake is driven by villages that were not initially connected to commercial parcel delivery services. The treatment effect is 10.6 percent among the roughly 85 percent of villages not previously connected to commercial parcel delivery but a relatively precise zero for villages with preexisting parcel delivery. On the production and local retail sides, we find no significant effects in either group of villages, confirming the earlier pooled results.[16] Turning to other potential sources of heterogeneity, we find that younger, richer households that are in closer proximity to the planned terminal and in more remote villages experience stronger uptake on the consumption side. For example, consumption uptake would close to double if average incomes were to double and primary earners were on average ten years younger. Somewhat surprisingly, we find no significant heterogeneity with respect to the years of education.

Spillovers.—We investigate the role of spillovers that could bias findings from the survey data. For example, if trade linkages with surrounding villages are an important driver of the local economy, then the comparison between treated and control villages could miss income or retail price effects. More simply, residents in control villages could use e-commerce terminals in a nearby treated village. To investigate these forces, we follow Miguel and Kremer (2004) and use variation in a village's exposure to other nearby treated villages after controlling for proximity to all villages (see online Appendix C). On the consumption side, we find evidence of positive spillovers from nearby terminals in other villages, as previewed above. In contrast, we find no evidence of cross-village spillovers on retail stores or on the production side. Consistent with the absence of income or price spillovers, we also confirm in microdata from the 2010 census that the fraction of village market access driven by trade with other nearby rural markets is minor (less than 3 percent).[17]

B. *Evidence from Firm Database*

We use the firm's internal transaction database to provide evidence on two questions that are outside the scope of the fieldwork.[18] First, to what extent are consumption and production responses to e-commerce access increasing beyond

[16] In line with the pooled results, online Appendix A reports some evidence that effects on product additions and stores sourcing online are stronger in villages without preexisting parcel delivery.

[17] Given how small villages are compared to cities, and that a small fraction of all villages participate in the program, GE effects on urban centers are unlikely in our setting.

[18] Online Appendix D also uses these data to investigate the representativeness of our RCT village sample and the timing/seasonality of the survey data collection.

our survey's 12-month time window? Second, are our survey data missing rare but highly successful tail events on the production side that could shift the average effect on local household incomes?

To answer these questions, we use the universe of transaction records from 5 provinces and about 12,000 villages that had been treated by April 2017 to estimate the following event study specification:

$$y_{vm} = \theta_v + \delta_m + \sum_{j=-3}^{24} \beta_j MonthsSinceEntry_{jvm} + \epsilon_{vm}, \tag{2}$$

where v indexes villages, δ_m is a set of month fixed effects between November 2015 and April 2017, and θ_v is a village fixed effect. Each observation in equation (2) is a village in a given month. The variable y_{vm} is one of four village-level monthly outcomes: number of buyers, number of purchase transactions, number of out-shipments, and total weight of out-shipments in kilograms. We create a balanced panel in the sense that each of the villages 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 spans terminal observations of up to 17 months pre-installation for villages connected in April 2017 and up to 28 months post-installation for the earliest villages connected by the program. A negative index j denotes the number of months prior to program entry. A positive j indexes the number of months since the program started operation, so β_0 is a measure of average outcomes for villages 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 to or beyond 24 months after program entry, so β_{24} captures average outcomes among villages that have been in operation for more than two years. Each of the β_0–β_{24} is estimated relative to the omitted category that is the period preceding program entry (zeros by construction since the program did not exist).

Figure 1 presents the event study plots for village-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. Program usage increases rapidly for about two to four months after opening, and then plateaus at around 85 buyers and 280 transactions per month per village. On the production side, we find evidence that the number and total weight of out-shipments increase smoothly over time after program entry and beyond the 12-month window covered in our survey data. The effect increases by roughly 50 percent when comparing the point estimate on the total weight of out-shipments 12 months post-entry to that more than 2 years post-entry. These results suggest that production-side adjustments take longer to fully materialize than our survey's one-year horizon. Despite this positive trend, the average monthly estimated effects at the village level remain small more than two years post implementation at around ten out-shipments with a combined weight of 30 kilograms.

Turning to the second question, our sampling of 38 households per village in the survey data collection may be insufficient to capture rare but very successful events on the production side. To investigate this issue, we use the universe of out-shipments depicted in Figure 1 and make the following assumptions to get an upper-bound estimate for these shipments' potential income creation in the local village economy: we assume (i) that the entire value of these shipments is local

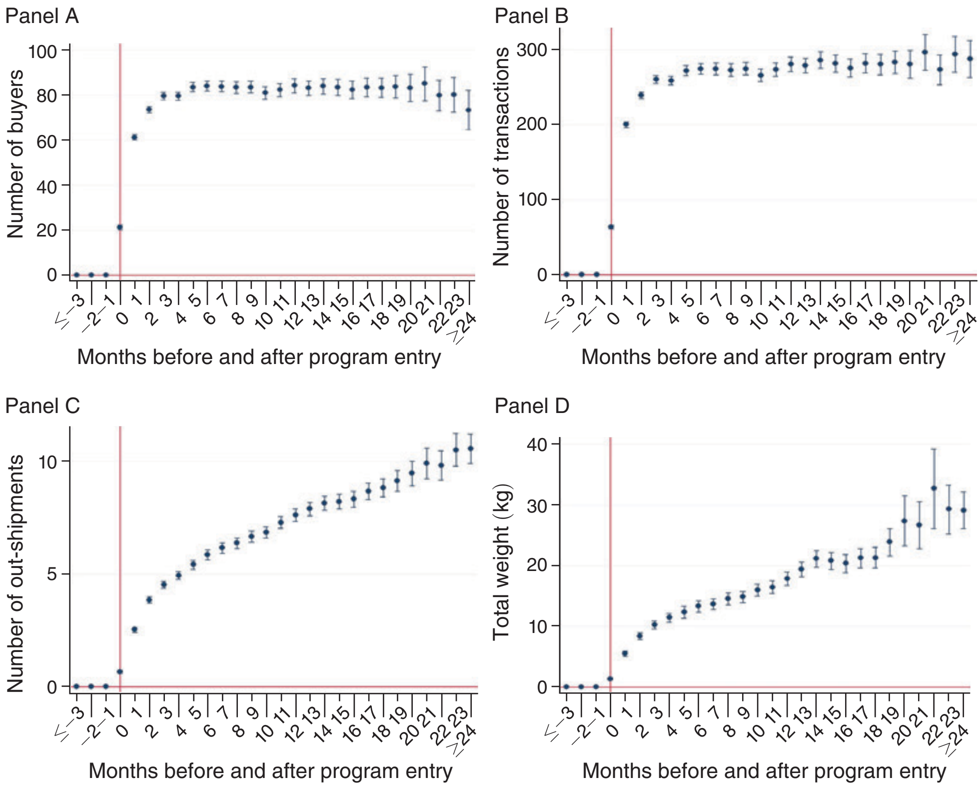

FIGURE 1. TIMELINE OF ADJUSTMENT: VILLAGE E-COMMERCE CONSUMPTION AND OUT-SHIPMENTS

Notes: Figure shows point estimates from a regression of depicted outcomes on months since program entry and village and month fixed effects. Outcomes are the number of buyers (panel A), number of transactions (panel B), number of out-shipments (panel C), and total weight of out-shipments (panel D) per village. The data are from the e-commerce firm's internal database and contain the universe of village purchase transactions from November 2015 to April 2017 and the universe of sales transactions from January 2016 to April 2017 in the five provinces of Anhui, Guangxi, Guizhou, Henan, and Yunnan (roughly 12,000 villages in total). The last point estimate of each plot pools months 24 to 28. The figure shows 95 percent confidence intervals based on standard errors that are clustered at the village level. See Section IIB for discussion.

value added and (ii) that the average value per kilogram of these shipments is as high as that of Chinese exports to the rest of the world.[19] Even 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. To conclude, this average longer-term effect—that we can estimate precisely in Figure 1 using the firm's transaction data—would still be consistent with the statistical zero results on incomes and the production side that we find using the RCT survey data after one year.[20]

[19] On average ¥66.50 per kilogram in 2015 and 2016 (World Integrated Trade Solution database).

[20] Related to this, much of the existing literature on information and communication technology in developing countries have estimated effects after relatively short periods. For example, Jensen (2007) documents significant effects of Indian cell phone towers on market prices and other outcomes within weeks post-installation. More recently, Hjort and Poulsen (2019) document effects of fast-speed internet on local employment and incomes in Africa that arise within 3–12 months post-installation.

III. Evaluation

In the final part, we interpret the program's observed effects through the lens of a simple theoretical framework. The most robust effect that we find is on the substitution of local households' retail expenditures to the new e-commerce shopping option. To quantify the cost of living implications consistent with these estimates, we follow a revealed-preference approach as in recent work by Atkin, Faber, and Gonzalez-Navarro (2018) and structure household preferences into three tiers: the upper tier is Cobb-Douglas over broad product groups $g \in G$ (durables and nondurables) in total consumption, the middle tier is CES across retailers $s \in S$ selling that product group (for example, local stores, market stalls, or the e-commerce option), and the final tier is across individual products within groups $b \in B_g$, which can be left unspecified (see online Appendix E for more details). The direct consumer gains from the arrival of the e-commerce option, measured as a percentage of initial household expenditure, can then be expressed as follows:

$$(3) \qquad \frac{Gains_h}{Initial\ Expenditure_h} = \prod_{g \in G} \left(\left(\sum_{s \in S_g^C} \phi_{gsh}^1 \right)^{\frac{1}{\sigma_g - 1}} \right)^{\alpha_{gh}} - 1,$$

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

To estimate this expression, we require information about the program's effect on $\sum_{s \in S_g^C} \phi_{gsh}^1$ and the parameters α_{gh} and σ_g. For the α_{gh}, we use our baseline data on household expenditure shares across product groups. For ex post expenditure shares on the new e-commerce option, we use the treatment effects among the 85 percent of villages without preexisting parcel delivery connections reported in Table 2. These villages experienced the removal of both logistical and transactional barriers to e-commerce trading. We include mean program usage among control villages in these treatment effects to account for program spillovers as discussed above.

We perform this welfare computation for two different groups of local households: first for the average sample household, for whom the average effect on the terminal share of total retail consumption is 1.6 percent, and second for households that report ever having used the terminal for consumption, for whom this effect is 14 percent. We also estimate price index effects separately for durable and nondurable consumption. And we report estimates both with and without reweighting households according to sampling weights. Finally, we calibrate σ_g using estimates from Atkin, Faber, and Gonzalez-Navarro (2018) for households in Mexico with incomes comparable to those of rural Chinese households in our survey ($\sigma_N = 3.87$ for nondurables and $\sigma_D = 3.85$ for durables).

Table 3 reports the estimation results. The average reduction in retail cost of living among households that experienced the lifting of both logistical and transactional barriers is 0.82 percent. This effect increases to 5.6 percent among the roughly 15 percent of households that ever used the new e-commerce option. These effects

Table 3—Average Effects on Household Welfare

	Unweighted (effects in sample)			Weighted (effects in village population)		
	Durables consumption	Nondurables consumption	Total retail consumption	Durables consumption	Nondurables consumption	Total retail consumption
Reduction in retail cost of living for all households	3.379%	0.481%	0.824%	2.962%	0.429%	0.73%
	(0.03)	(0.003)	(0.005)	(0.03)	(0.003)	(0.005)
Reduction in retail cost of living among users	19.884%	3.806%	5.597%	16.637%	3.217%	4.722%
	(0.221)	(0.028)	(0.034)	(0.224)	(0.025)	(0.032)

Notes: Table reports average household gains in terms of percentage point reductions in retail cost of living for different consumption categories and groups of households. Estimates are based on equation (3) using treatment effects on household substitution into the new e-commerce option. The left panel reports unweighted results, and the right panel adjusts the weight of each household using sampling weights. Standard errors are bootstrapped across 1,000 iterations, taking into account that the treatment effects are point estimates. See Section III for discussion.

are slightly lower at 0.73 and 4.7 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: 3 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.[21]

Finally, to investigate the distribution of these gains, we use treatment effects from the joint heterogeneity specification in Table 2, panel G. We estimate this specification with the dependent variable being the household expenditure share on the new e-commerce option for either durables or nondurables. For each sample household in treatment villages without preexisting parcel delivery, we then compute a fitted value of the effect on $\sum_{s \in S_g^C} \phi_{gsh}^{t1}$, based on the primary earner's age, income per capita, residential distance to the planned terminal, and distance to the nearest township center (remoteness), included jointly. Figure 2 shows these graphs. Ranking households along each of these dimensions, we find more than fourfold differences in the price index effect within the sample. For example, the average rural household with a 25-year-old primary earner experiences a reduction in retail cost of living of about 1.5 percent (without conditioning on uptake), which drops below 1 percent past the age of 40 and close to 0 past the age of 60.

Overall, our findings suggest that the welfare gains from e-commerce trading access are limited to certain groups of rural households and particular markets rather than being broad-based. First, we show that the income and production-side effects that have been the focus of the existing literature on e-commerce villages are not representative of the countryside, even when focusing on a sample of rural markets in the RCT that were chosen by the firm for successful e-commerce expansion. Second, we find strong heterogeneity in the consumer gains from e-commerce across villages and households within them. In this light, we hope this work can inspire additional research to investigate what types of local factors or complementary

[21] We also evaluate robustness to alternative σ_g. Assuming $\sigma_N = 2.87$ and $\sigma_D = 2.85$ yields larger gains (a 1.27 percent reduction in retail cost of living on average and 8.74 percent among users). Assuming $\sigma_N = 4.87$ and $\sigma_D = 4.85$ yields slightly smaller effects (0.61 and 4.12 percent, respectively).

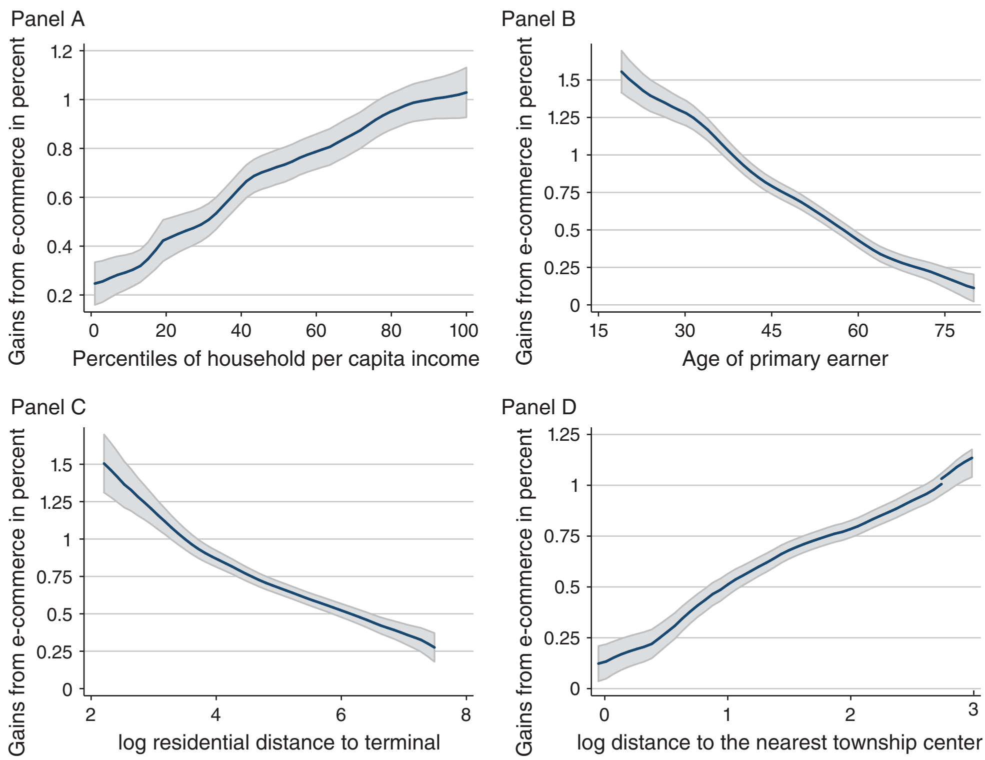

Figure 2. Heterogeneity of Gains from E-Commerce

Notes: Figure shows predicted average gains (users and nonusers) in terms of percentage point reductions in household retail cost of living as a function of household per capita income (panel A), age of primary earner (panel B), residential distance to terminal (panel C), and distance to the nearest township center (panel D). Predictions are based on treatment effects from Table 2, panel G. The figure depicts 95 percent confidence intervals that are based on clustering standard errors at the village level. See Section III for discussion.

interventions allow rural markets to reap the gains from trade through e-commerce for both producers and consumers.

REFERENCES

AliResearch. 2018. *China's Taobao Villages*. Hangzhou: Alibaba Group Research Department.

Asher, Sam, and Paul Novosad. 2020. "Rural Roads and Local Economic Development." *American Economic Review* 110 (3): 797–823.

Atkin, David, Benjamin Faber, and Marco Gonzalez-Navarro. 2018. "Retail Globalization and Household Welfare: Evidence from Mexico." *Journal of Political Economy* 126 (1): 1–73.

Brynjolfsson, Erik, Yu Hu, and Michael D. Smith. 2003. "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers." *Management Science* 49 (11): 1580–96.

Casaburi, Lorenzo, Rachel Glennerster, and Tavneet Suri. 2013. "Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone." https://scholar.harvard.edu/files/lorenzocasaburi/files/casaburi_glennester_suri_roads.pdf.

Chapman, Robert, and Tom Slaymaker. 2002. "ICTs and Rural Development: Review of the Literature, Current Interventions and Opportunities for Action." Overseas Development Institute Working Paper 192.

China Internet Network Information Center. 2015a. *Research Report on China's Online Shopping Market in 2014*. Beijing: China Internet Network Information Center.

China Internet Network Information Center. 2015b. *The 36th Statistical Report on Internet Development in China*. Beijing: China Internet Network Information Center.

Connor, Neil. 2016. "Inside China's Tech Villages Where Farmers Have Replaced Communism with E-Commerce." *Telegraph*, November 5. https://www.telegraph.co.uk/news/2016/11/05/inside-chinas-tech-villages-where-farmers-have-replaced-communis/.

Couture, Victor, Benjamin Faber, Yizhen Gu, and Lizhi Liu. 2017. "E-Commerce Integration and Economic Development: Evidence from China." AEA RCT Registry. June 1. https://doi.org/10.1257/rct.1582-5.0.

Couture, Victor, Benjamin Faber, Yizhen Gu, and Lizhi Liu. 2021. "Replication Data for: Connecting the Countryside via E-Commerce: Evidence from China." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E117506V1.

Dobbs, Richard, Yougang Chen, Gordon Orr, James Manyika, Michael Chui, and Elsie Chang. 2013. *China's E-Tail Revolution: Online Shopping as a Catalyst for Growth*. New York City: McKinsey Global Institute.

Dolfen, Paul, Liran Einav, Peter J. Klenow, Benjamin Klopack, Jonathan D. Levin, Larry Levin, and Wayne Best. 2017. "Assessing the Gains from E-Commerce." Unpublished.

Egypt Ministry of Communications and Information Technology. 2016. *Egypt's National E-Commerce Strategy*. Giza Governorate: Ministry of Communications and Information Technology.

Fan, Jingting, Lixin Tang, Weiming Zhu, and Ben Zou. 2018. "The Alibaba Effect: Spatial Consumption Inequality and the Welfare Gains from E-Commerce." *Journal of International Economics* 114: 203–20.

Federal Reserve Bank of St. Louis. 2016. "E-Commerce Retail Sales as a Percent of Total Sales (ECOMPCTSA)." Federal Reserve Economic Data. https://fred.stlouisfed.org/series/ECOMPCTSA.

Forman, Chris, Avi Goldfarb, and Shane Greenstein. 2012. "The Internet and Local Wages: A Puzzle." *American Economic Review* 102 (1): 556–75.

Freedman, Josh. 2017. "Once Poverty-Stricken, China's "Taobao Villages" Have Found a Lifeline Making Trinkets for the Internet." *Quartz*, February 12. https://qz.com/899922/once-poverty-stricken-chinas-taobao-villages-have-found-a-lifeline-making-trinkets-for-the-internet/.

Goldmanis, Maris, Ali Hortaçsu, Chad Syverson, and Önsel Emre. 2010. "E-Commerce and the Market Structure of Retail Industries." *Economic Journal* 120 (545): 651–82.

Goyal, Aparajita. 2010. "Information, Direct Access to Farmers, and Rural Market Performance in Central India." *American Economic Journal: Applied Economics* 2 (3): 22–45.

Hjort, Jonas, and Jonas Poulsen. 2019. "The Arrival of Fast Internet and Employment in Africa." *American Economic Review* 109 (3): 1032–79.

India Ministry of Electronics and Information Technology. 2016. *Digital India: Driving E-Commerce in Rural and Semi-Urban India*. New Delhi: India Ministry of Electronics and Information Technology.

Jensen, Robert. 2007. "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector." *Quarterly Journal of Economics* 122 (3): 879–924.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Luo, Xubei, Yue Wang, and Xiaobo Zhang. 2019. "E-Commerce Development and Household Consumption Growth in China." World Bank Policy Research Working Paper 8810.

Luo, Xubei, and Chivu Niu. 2019. "E-Commerce Participation and Household Income Growth in Taobao Villages." World Bank Poverty and Equity Global Practice Working Paper 198.

Miguel, Edward, and Michael Kremer. 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*. Allen, Texas: PSFweb.

Prime Minister of Vietnam. 2016. *Vietnam's E-Commerce Development Masterplan*. Ha Noi: Vietnam's Office of the Prime Minister.

Statista. 2016. "Online Retail Statistics for China." Market Statistics.

UNCTAD. 2016. "E-Commerce Opens New Opportunities for Developing Countries." http://unctad14.org/Documents/U14_IN_2016_011_Ecommerce_En.pdf.

van de Walle, Dominique. 2009. "Impact Evaluation of Rural Road Projects." *Journal of Development Effectiveness* 1 (1): 15–36.

World Bank. 2016. *World Development Report 2016: Digital Dividends*. Washington, DC: World Bank.

Weller, Chris. 2017. "Chinese 'Taobao Villages' Are Turning Poor Communities into Huge Online Retail Hubs." *Business Insider*, February 27. https://www.businessinsider.com/chinese-taobao-villages-escape-poverty-2017-2.