Trading through Platforms: Evidence from AliExpress*

Richard Baldwin † CEPR

Edoardo Chiarotti ‡ Graduate Institute

Daria Taglioni § World Bank

January 30th, 2021

Abstract

This paper strives to contribute three things to the literature. The first is a line-sketch theory model that puts consumers into the value chain. It emphasizes the value-creating aspect of superior matching between consumers preferences and the varieties when they can purchase online (and thus have access to a broader range of varieties than is available locally). We indirectly test this by estimating the impact that AliExpress has on exports. AliExpress data is unique in that only Chinese firms can sell and only non-Chinese can buy on the platform. Third, we provide evidence that the introduction of such platforms can support export of domestic value added.

JEL Classification: F10, F14, L81

Keywords: Platforms, AliExpress, Trade

*We thank participants at GLOBECOS - Global Community of Social Sciences, UIC Barcelona seminar on Collaborative Economy in the Digital Environment, Varsaw International Economic Meeting, and Global Economic Policy Group 2020. This working paper is funded by the Swiss National Science Foundation (Grant N. 159346). The views expressed herein are those of the authors only and not reflect those of Center for Trade and Economic Integration, Center for Economic Policy Research, the World Bank, or any other institution with which the authors are affiliated. Errors are our own.

†Email: rbaldwin@cepr.org.

‡Email: edoardo.chiarotti@graduateinstitute.ch.

§Email: dtaglioni@worldbank.org.

1 Introduction

A recent survey defined global value chains as arising when "production stages for individual goods are broken apart and scattered across countries" (Johnson, 2018). This is natural since most economic analyses of global value chains focus on fractionalization and dispersion of production (Grossman and Rossi-Hansberg, 2008; Baldwin and Venables, 2013, etc.). Put differently, the baseline conceptualization of global value chains (GVs) focuses on business-to-business (B2B) relationships; the role of the consumer is set aside for clarity's sake.

Traditionally, putting consumers aside (in the theory sense) has been viewed as an innocent simplification. The so-called "Platform Revolution", however, has rob this assumption of some of its innocence. Platforms are no longer mere distributional channels - they are inserting consumers into the value chain. They are intermediating trade in ways that allow new forms of value creation based on matching production and consumers preference with a higher degree of resolution than is typically possible via brick-and-mortar retail outlets. In this sense, platforms allow consumers to join value chain by transmitting their preferences directly to producers. As Parker et al. (2016), put it: "The platform concept is fundamentally simple: create a place where producers and consumers come together in interactions that create values for both parties."

Since this distinctive role for consumers may be unfamiliar, it is worth drawing an analogy with a distinction that is widely recognized, namely the difference between producer-led and buyer-led value chains. Producer-led chains, for example those observed in motor vehicles and consumer electronics, focus purely on the production process. Examples of buyer-led chains including "fast fashion" retailers like Zara and supermarket-led chains that sell processed food under their own label.

A key difference between the two types is the role of asymmetric information. The key advantage that Zara has is its knowledge of exactly what young consumers want to buy this week. In the fast-fashion example, there is a firm intermediating between the consumers and the producers, but knowledge of preferred varieties is the key to success. One can think of online e-commerce platforms as removing the intermediate role of retailers; they allow consumers to directly communicate their preferences to producers. Just as Zara is creating value by directing suppliers to make the varieties that their customers actually want, the platforms are creating value by reducing information asymmetries in the market. The mechanism we stress in this paper predicts that the opening of platforms leads to more trade since it gives consumers access to a wider range of varieties to choose from. It also predicts that online trade does not replace traditional trade since shipping and handling costs are lower per kilo for traditional trade. In a sense we view the online platforms as more akin to high-end boutiques than to bargain-based discounters. People shop in boutiques, despite the higher prices, since the boutiques carry varieties that are not available in the mass-market retailers.

1.1 Literature Review

The recent years have seen the rise of platform firms like Amazon and Ebay as the new market place between global manufacturers and consumers (World Bank, 2020). A new academic literature has analysed how such platforms drastically changed the way countries, firms and people trade and participate to global value chains (GVCs). Generally, the literature points out how platforms decrease frictions to trade (e.g. Hortaçsu et al., 2009; Lendle et al., 2016; Brynjolfsson et al., 2019; Li et al., 2019), but often at the cost of increasing market concentration (e.g. Duch-Brown and Martens, 2014; Bai et al., 2018; Chen and Wu, 2020) and excessive bargaining power by platforms' owners (Zhu and Liu, 2018). While these findings are consolidated for developed countries, the lack of data makes it more difficult to study whether platforms played a similar role also in emerging economies.[1]

The literature on trade, e-commerce and platforms includes Hortaçsu et al. (2009) and Lendle et al. (2016). These papers document the "death of distance" result that geographic distances appears to be less important for online sales than it is for traditional trade. Hortaçsu et al. (2009), Goldmanis et al. (2010) and Lieber and Syverson (2012) argue that these results suggest that the main benefit of platforms as a trade facilitator comes via its impact on search costs. This axis of investigation has been connected to the suggestions Chaney (2014) and Allen (2011) that the very large impact of geographical distance on trade volumes mostly reflects information frictions. The mechanism we stress also differs from the one considered by Goldmanis et al. (2010), who look at how platforms affect the evolution of the retail sector. They show that lower search costs shift shares from high cost to low

[1]As a rare example, Couture et al. (2018) show that, while investing in logistical programmes to ship goods to Chinese villages increase the use of e-commerce in these villages, the households befeniting most are the ones in better positions to receive e-commerce benefits, i.e. younger and richer.

cost stores.

Our paper is also related to the strand of literature studying how platforms can reduce frictions to trade. In general, in the offline world factors like weak local institutions and lack of trust in local suppliers hinder trade (Anderson and Marcouiller, 2002; Guiso et al., 2009; Allen, 2014; Steinwender, 2014; Dasgupta and Mondria, 2018). Differently, in the online world technology may help to reduce matching frictions and therefore price dispersion (Brynjolfsson and Smith, 2000; Brown and Goolsbee, 2002; Ghose and Yao, 2011; Overby and Forman, 2015; Einav et al., 2016; Carballo et al., 2016; Goldfarb and Tucker, 2019). Specifically, platforms reduce search costs and help define precise contracts via tools like consumer rating (Hortaçsu et al., 2009; Goldmanis et al., 2010; Lewis, 2011; Lendle et al., 2016). The literature documents that this is the case for different types of platforms, such as car-sharing, lodging and e-commerce platforms (Lam and Liu, 2017; Liu et al., 2018; Farronato and Fradkin, 2018; Fan et al., 2018; Hui, 2020). Some authors have also examined the cases of Alibaba's domestic and cross-border trade platforms, namely Taobao and AliExpress. In particular, this literature reports that, on these platforms, sellers' rating-measured reputation drives their performance and implies concentration of superstar exporters (Fan et al., 2016; Li et al., 2016; Chen and Wu, 2020).

Our paper also contributes to the more general literature addressing the impact of information and communication technology on international trade and global value chains. The majority of papers finds a positive correlation between the use of ITC and trade (Freund and Weinhold, 2004; Demirkan et al., 2009; Vemuri and Siddiqi, 2009; Márquez-Ramos and Martínez-Zarzoso, 2010; Timmis, 2012; Mattes et al., 2012; Freund et al., 2019). This channel may work also in emerging economies (Baliamoune, 2002; Clarke and Wallsten, 2006; Hjort and Poulsen, 2019; Fernandes et al., 2019), but it can lose strength in countries with lower income per capita (Portugal-Perez and Wilson, 2012).

1.2 Prima Facie Evidence

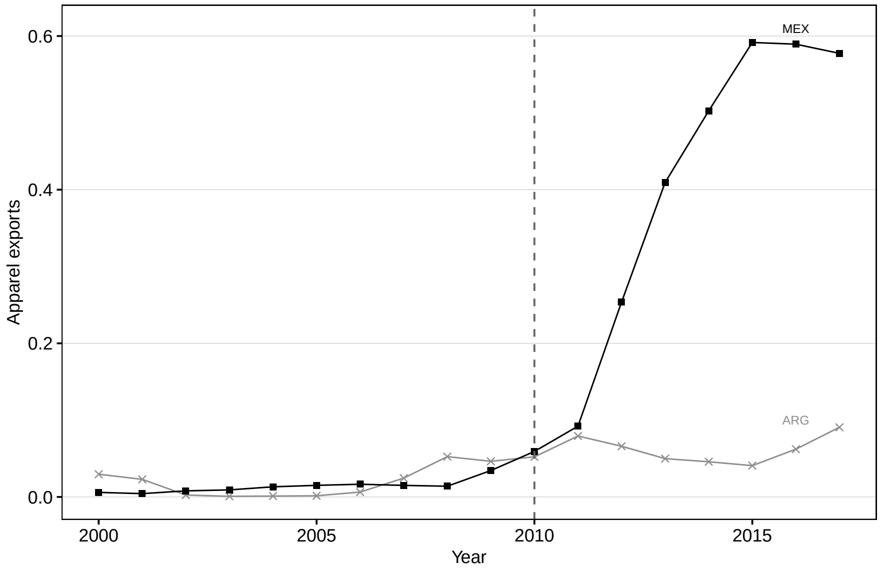
Our model predicts that the opening of platforms leads to more trade since it gives consumers access to a wider range of varieties to choose from, but it predicts that online trade does not replace traditional trade since the shipping technologies used by traditional trade are cheaper due to non-convexities and the large volume of shipments to retail outlets. Data from the online platform AliExpress allows us to address at least one aspect of this. The online-retail platform, created by AliBaba in 2010, had some very particular features. First of all, it is large. In 2013, it had 1.1 million active sellers, a traffic flow of 3.8 million consumers each day, and a sales volume of nearly 113 billion orders worth a total of $20 billion. Second, it is global. Today, AliExpress has over 50 million product listings, consumers in over 220 countries and a global market share that exceeds Amazon and eBay (Chen and Wu, 2020). Third, and mostly interestingly, the platform's rules dictate that only Chinese firms can sell goods on the platform and only non-Chinese consumers can buy on the platform. As a result, all the web traffic on AliExpress.com outside of China is carried over by consumers only and we can exploit the geographical distribution of AliExpress visitors to predict the patterns of Chinese exports.

In principle, after 2010, we expect to see larger increases in Chinese exports towards those trade partners with larger shares of AliExpress visitors, and specifically in those sectors that can be traded on AliExpress. This is indeed what we find with our empirical work, but a quick plot suggests that the phenomenon is significant. Figure 1 shows that this is the case if we compare Chinese exports in the apparel sector - which is largely traded on AliExpress (Chen and Wu, 2020) - towards Argentina and Mexico. Mexico has a share of AliExpress visitors that is approximately three times larger than Argentina. As you can see, such exports diverge significantly between Mexico and Argentina after the introduction of AliExpress in 2010.

A similar trend can be observed when we extend this exercise to the 42 trade partners for which we have a measure of the share of AliExpress visitors. Figure 2 reports the log change in Chinese exports before and after 2010 in the apparel sector for the 42 trade partners in the sample (y axis), ordered by the natural logarithm of visitors on AliExpress.com (x axis). As you can see, post-2010 increase in Chinese exports are larger for trade partners with larger shares of AliExpress visitors.

Obviously the jump up in trade could be consistent with many mechanism, but since the figures we show are total trade, it seems that the new sales channel is adding to, and not just replacing the traditional channel (via wholesalers and retailers). This suggest that it is there is some tradeoff that is keeping consumers buying from online and in-person.

Figure 1: Chinese Exports in the Apparel Sector - Mexico vs Argentina



Notes: Data for Chinese exports in the apparel sector for Mexico and Argentina are from Comtrade. Data on the number of visitors on AliExpress.com are from Alexa.

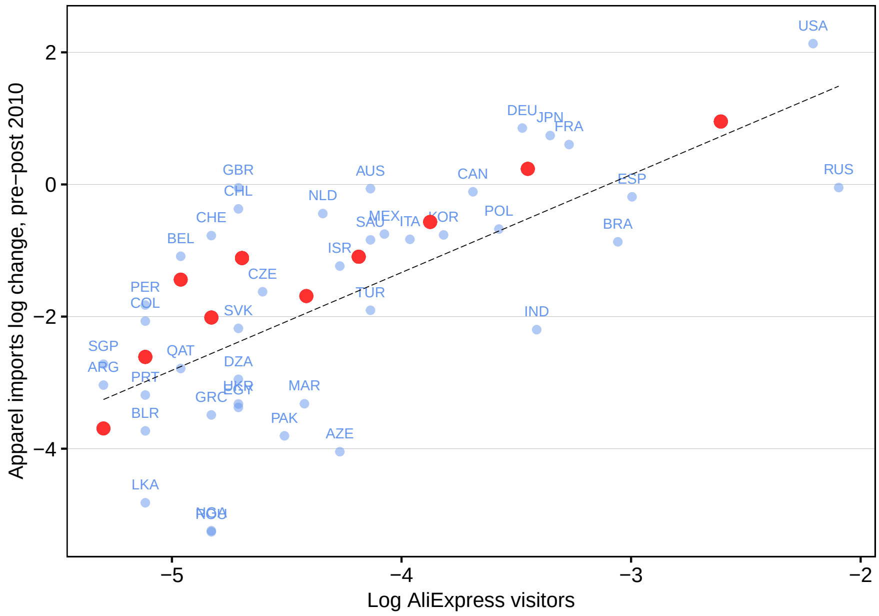
1.3 Contributions of the paper

This paper contributes to the literature by stressing a novel mechanism by which e-commerce is adding value by allowing consumers to join the value chain directly. The paper also contributes to trade and e-commerce literature by studying whether platforms have helped firms in emerging economies to trade more. To do so, we focus on the case of AliExpress, the export-focused, online-retail platform created by Alibaba in 2010 with the aim of helping small and medium Chinese enterprises to enter the export market and sell their products abroad.

We pick AliExpress because of a peculiar characteristic of this platform's rules, by which only Chinese firms can sell goods on AliExpress and only non-Chinese consumers can buy goods on AliExpress. As a result, all the web traffic on AliExpress.com outside of China is carried over by consumers only and we can exploit the geographical distribution of AliExpress visitors to predict the patterns of Chinese exports. In principle, after 2010, we expect to see larger increases in Chinese exports towards those trade partners with larger shares of AliExpress visitors, and specifically in those sectors that can be traded on AliExpress.

In our baseline specification, we find that when moving from a trade partner at the

Figure 2: Chinese Exports in the Apparel Sector - All Trade Partners



Notes: Data for Chinese exports in the apparel sector are from Comtrade. Data on the number of visitors on AliExpress.com are from Alexa. The log-percentage changes before and after 2010 are from authors' calculation. The light blue points show the country averages, while the red points show quantile averages for ten quantile bins of the AliExpress variable.

10th to 90th percentile of the distribution of the share of AliExpress visitors, the post-2010 increase in Chinese exports in the apparel sector is larger. This difference (in differences) amounts to about 32 million dollars when the US is included in our sample. After testing the robustness of our results, we extend our analysis to multiple export sectors. We consider the HS-2 classification and we define a bivariate variable which equals one if a sector can be traded on AliExpress and zero otherwise. We interact this variable with our DID term and we find that the AliExpress effect is larger for sectors that can be traded on AliExpress. In the second part of the paper, we assess whether the AliExpress-induced increase in exports came to the advantage of Chinese (rather than foreign) value added. To test this hypothesis, we rely on both an index proposed by Rauch (1999), measuring the level of products' differentiation in a sector, and data on Chinese exports in domestic value added from the TiVA database. Respectively, we find that the AliExpress effect is larger for differentiated products and that our results are confirmed when we use Chinese exports in domestic value added as our dependent variable.

1.4 Plan of the paper

The rest of the paper is organised as follows. Section 2 lays down a simple theoretical framework modelling the participation of consumers in the global value chain. Section 3 describes the data used. Section 4 reports the baseline results. Section 5 extends the analysis to multiple sectors. Section 6 presents the GVC analysis. Section 7 concludes.

2 A simple model of consumer purchases from retail versus online platforms

What are consumers doing when they buy via online platforms and how does this affect trade? One answer is that online purchasing affects trade by reducing transaction costs that separate suppliers and buyers. Search costs are usually the focus of this line of thinking which is the implicit mechanism of action in the trade and e-commerce papers that focus on the "death of distance" aspects, such as Lendle et al. (2016).

In this section we suggest an alternative approach - one that focuses on the fact that many consumers buy both online and in brick-and-mortar stores near them. The narrative of our model turns on the simple fact that a vastly greater range of varieties are available online than locally (say in retail outlets). Consumers in our model create value by providing better matching between varieties produced and varieties purchased.

To set the stage, we start with the classic Dixit-Stiglitz-Krugman trade model where varieties are doubly symmetric. They are symmetric on the supply side since the cost function is identical for every variety and firm in every nation. They are also symmetric on the demand side since the representative consumer's demand function is identical for each variety.

Firms must choose which variety they produce, but the choice is a distinction without a difference for the simple reason that the demand and degree of competition are identical for every variety. As is well known, the monopolistically competitive equilibrium involves average cost pricing. To be specific, we call this price p' and note that it is identical for all varieties in all markets (given the lack of trade costs

in the baseline model).
In this model, consumers face no choice among varieties (they find it optimal to buy some of each variety on offer), so we must add some elements to the model to bring behaviors in line with the model's narrative. In this frictionless world, a variety that is produced anywhere is sold everywhere. What is missing is a mechanism for limiting local choices.
To distinguish our approach from other approaches, we assume search costs are zero for all agents in all settings. Search cost, in other words, are not part of the way platforms affect outcomes in our model.
One way to limit the range of varieties available locally is to impose retailers between consumers and producers and then to impose per-variety costs on the retailers so that retailers do not carry an infinite number of varieties. Note that this approach differs from the standard approach to limiting varieties. In Arkolakis et al (2011), for example, it is producers - not retailers - who choose the range of varieties on offer to consumers. Our retail-based approach has the advantage that it permits a natural way of allowing consumers to have the option of buying locally and/or buying online - once the model allows platforms.
Ignoring the possibility of online sales for the moment, consider how the frictionless-trade equilibrium is altered by the introduction of retailers and per-variety costs (which we call shelf-space costs). While consumer would buy some of each variety on the shelves, retailers do not find it profitable to stock an infinite range of varieties. The reason is that an infinite range of goods would result in an infinitely small purchase per variety, and in this case, the operating profit per variety would be insufficient to cover the finite per-variety shelf-space cost.
In equilibrium, retailers, who faces a sourcing cost of p' per unit for every variety, would expand the range of varieties up to the point where the sales per variety yielded an operating profit per variety that just covered the per-variety shelving cost. Given the maintained assumption of complete symmetry on the demand side (across consumers and varieties), retailers are indifferent as to which varieties they stock. All that matters at this point is number of varieties. To give it a label, we call this number, n'. To recap, n' is number of varieties available in each local market and they are all priced at p' plus the average (per unit) shelfing cost, which we call b'. Since there is nothing coordinating the choice of varieties by retailers in different local markets, and there is an uncountable infinity of varieties produced, we can presume that each retailer chooses a distinct set of varieties.

How many varieties would be produced worldwide in this setting? Given the simplicity of the supply side, the answer lies in the number of varieties demanded. Aggregating across all local retailers, each of which orders n' varieties, the total number of varieties demanded is m times n', where m is the exogenously given number of local markets. To recap, the measure of varieties produced worldwide is mn'; each consumer only has access to n' varieties.

What would happen if consumers were allowed to buy directly from all producers via a platform? Given the love-of-variety preferences, they would buy all mn' varieties produced directly from producers, paying p' per unit. This means that they would buy nothing from retailers since direct purchases eliminate the need to cover the average shelf-cost (p' is less than $p' + b'$). Clearly additional elements are needed in the model to bring it into line with the narrative. There are two very distinct reasons that consumer buy online. First, they buy some of the varieties are not available locally. Second, they buy the locally available varieties online since purchasing on the platforms allow them to circumvent the average shelf costs, b', on those varieties.

Complete displacement of retailers may occur in the future (and indeed in market such as airline tickets it has already occurred), but it does not line up with the narrative of our model. To get consumers buying both online and in person, we need online-buying to confront consumers with disadvantages that balance the advantage of greater variety. To this end, we add an additional element to the model. We assume that the per-unit trade costs are higher for consumers purchasing directly from firms online than it is for retail stores. The notion here is that e-commerce typically involves shipping technologies (say, postal packages) that are less efficient per kilo, or per dollar of value than the technologies the retail stores uses (say containers). For clarity, we continue to assume zero trade cost for retailers, but add t' for goods bought online.

With this additional element in the model, consumers face a trade-off when purchasing online. On one hand, the platforms give them access to varieties that are not available locally. On the other hand, the landed price of online-bought goods is $p' + t'$, while the price of local varieties is $p' + b'$. If $t' > b'$, then consumers will buy all the locally available varieties at the store, and all the other varieties online. A moment's reflection shows that the platforms have created value. They introduce a previously unavailable channel, or technology, for selling/buying. The new technology has higher variable costs (b') but lower fixed costs (no shelf costs), so it does not dominate, or replace the old technology. What it does is create a gain from trade based on the consumption of a wider variety of goods. To see, note that the ideal price index for the CES preferences we are working with, namely $P = (n'(p'+b')^{1-\sigma} + n''(p'+t')^{1-\sigma})^{1/(1-\sigma)}$ where n'' is the number of varieties bought online, falls as n'' rises:

$$\frac{dP}{P} = \frac{(1-\sigma)(p'+t')^{1-\sigma}}{n'(p'+b')^{1-\sigma} + n''(p'+t')^{1-\sigma}} \frac{dn''}{n''}$$

Since $\sigma > 1$ as usual in the baseline model, the price index falls as n'' rises (despite the higher price of varieties bought online. This is the source of the gains from trade. The model so far still does not conform with our narrative since it is the technology of the platform that is creating the additional value - not the choices made by consumers. Platforms are allowing consumers to participate in the value chain directly, but it is not really the consumers who are adding the value. It is the novel trade technology and thus more in line with the "transaction cost" approach than our "consumption matching" approach.

What we need is elements in the model that does two things: 1) forces consumers to choose among varieties, 2) makes the choice matter in the sense of changing the outcome of matching varieties and consumers. To force a choice on consumers, we add a positive, finite per-variety "closet space" cost for consumers so that it is never optimal for them to buy an uncountable infinity of varieties. They have to choose which varieties to buy. To make the choice consequential, we make consumers heterogeneous in that they prefer some varieties to others. A simple way to add the latter without deviating too far from the baseline model is the quality approach of Baldwin and Harrigan (2011).

They consider CES preferences where consumption quantities are weighted by a quality factor, i.e. $U = (\int_{i\in\Theta}(c_i q_i)^{1-1/\sigma} di)1/(1-1/\sigma)$. Here the q's are the subjective quality as perceived by the consumer and it is convenient to think of them as random variables whose realizations are known to the consumer but not to firms are retailers. The corresponding, consumer-specific expenditure functions (and thus demand functions) are:

$$p_i c_i = \frac{p_i/q_i}{P}^{1-\sigma} E; \qquad P \equiv (\int_{j\in\Theta} (p_j/q_j)^{1-\sigma} dj)^{1-\sigma}$$

where E is the expenditure. For simplicity, we assume that the preference differences across consumers reproduce the homogenous consumer case considered above once demands are aggregate across all consumers. Thus from the perspective of retailers and firms, the demand functions they face are unaltered, but from the perspective of consumers, some varieties are better than others. This imbues the choice of varieties with economic meaning.

Which varieties do consumer choose? To set up the problem, we start with only local purchases being possible, so the price of all varieties is $p' + b'$.

The structure of the problem is very similar - in a purely mathematical sense - to the problem of which firms export in the Melitz model. If we look at the utility index:

$$U^{(1-1/\sigma)} = \int_{i\in\Theta} (c_i q_i)^{1-1/\sigma} di$$

and plug in the optimal consumption level given prices and expenditure, we find that the utility provided by consuming variety-j is:

$$p_i q_i^{1-\sigma} B, \qquad B \equiv \frac{E}{P}$$

where P is the usual CES price index for all the varieties that the consumer consumers. This is the utility benefit of buying variety i; the cost of buying one more variety is the per-variety, closet-space cost which we label as h'. If we choose units to normal all the prices to unity, then the consumer will buy all varieties where the utility benefit exceeds the closet cost. The threshold level of perceived quality, q', is thus defined as:

$$q'^{\sigma-1} B = h'$$

If we order the varieties for a particular consumer so the q's are decreasing, then the consumer buys all varieties with $q_i < q'$. The final element is the distribution of consumers' preferences. We assume there is a continuum of consumers whose preferences are different but in aggregate they general preferences for varieties that are identical to what they would be if every consumer had the standard CES preferences. There are several assumptions on the distributions of q's that would generate this. One extremely simple one would have all consumers to have only one favored variety (i.e. $q_i > 1, q_j = 1, i \neq j$) and then to associate the consumer's index with the varieties index. The market demand that arises from integrating over all the consumers would produce perfectly symmetric demands for each variety from the perspective of firms and retailers.

What is the equilibrium with heterogenous consumers? Before the platform opens, consumers will buy varieties locally up to the point where the marginal utility from an extra variety equals the marginal closet cost. Individual consumer will, naturally, priorities varieties that - according to their idiosyncratic preferences - are higher quality. If closet costs are low enough, or the consumers favor only few varieties, then the equilibrium will involve the retail purchase of some varieties that are favored, and some that are not.

When the platform opens, a typical consumer will have access to a wider range of varieties and thus will select more varieties that are favored. As long as the marginal closet costs are flat, there will be no substitution of online varieties for local varieties so the volume of trade will rise. The key point is why it rises.

With heterogenous preferences, the platforms allow better matching between producers and consumers. This will result in a lower consumer-specific, quality-adjusted price index.

Stepping back, it is clear that platforms are playing a very different role in this model than they do in the standard approach. Consumers are not buying online because the with-costs prices are lower. They are not buying online because the search costs are lower. They are buying online since the platforms given them access to varieties that they cannot find in their local stores - varieties for which they are happy to pay a somewhat higher price.

Consumers in this model are adding value. The goods they buy online are worth more in terms of utility, i.e. have higher value, than the goods they buy locally - even taking account of the higher trade costs. Since the online purchases will change Chinese production patterns, the platforms - by letting consumers choose exactly the correct variety - are generating higher value added. Better matching is creating the extra value.

2.1 Testable hypotheses

Our model predicts that the opening of platforms leads to more trade. That is the most obvious testable hypothesis. Closely related is the model prediction that online purchasing does not displace traditional trade. That is, the consumers will buy both online and retail. The net result is that the bilateral trade flows should increase. A refinement of the trade-volume hypothesis is that there is no displacement effect in the sense that normal trade (mediated by retailers and wholesalers) does not diminish even as e-commerce rises.

What would be the empirical implications if the assumption about online sales being more expense were wrong? What should we see in the data if indeed platform sales reduced transactions costs? Since consumers are free to chose online or retail, we should see an important displacement effect. Indeed in the purely domestic market recent years have seen just such displacement.
To distinguish these effects, we can look at the time pattern of the impact of platforms on bilateral trade. If platforms are a cheaper way of trading, we should see a graduate rise in the importance of the platformâ€™s presence. If, by contrast, the platforms are just allowing consumers to pay a premium of preferred varieties, there need be no systemic rise in the impact of platforms on bilateral trade.

3 Data

3.1 Apparel Sector and Country Variables

For our baseline empirical analysis, we rely on a panel dataset on 42 trade partners of China over 2000-2018. Table 1 reports summary statistics for the main variables in our sample. The dependent variable of interest is Chinese gross exports in the apparel sector (HS-2 61) from Comtrade. We clean this data by winsorization below and above the, respectively, 1st and 99th percentiles. The measure on the number of AliExpress visitors by country is provided by Alexa and reports the average number of visitors on AliExpress.com over the months of October, November and December 2019, expressed as a percentage of total Worldwide visitor. Our controls include internet penetration, which is the percentage of individuals using the internet, provided by the International Telecommunication Union. We also control for various gravity variables, such as distance from China (source: CEPII), and macroeconomic indicators, such as GDP, GDP per capita and population (source: World Bank).

3.2 AliExpress

We pick AliExpress because, by the company's rules, only Chinese firms can sell products on AliExpress, and only non-Chinese consumers can buy them.[2] We can

[2] This was the case until 2019, when AliExpress started to allow also non-Chinese firms to sell products on its platform

Table 1

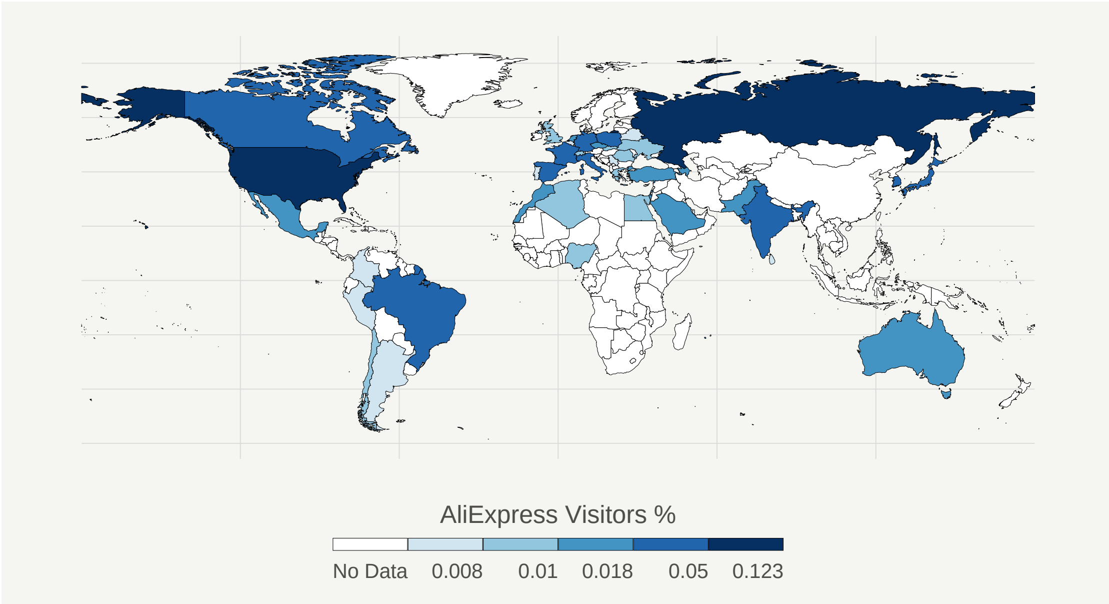
Statistic	N	Mean	St. Dev.	Min	Max
Export (Bil. US dollars)	812	1.02	2.43	0.0000	16.93
Export Winsorised (Bil. US dollars)	812	1.00	2.36	0.0001	14.64
AliExpress Visitors (%)	812	0.02	0.02	0.005	0.12
Distance (Tho. Km)	811	8.58	4.05	0.96	19.30
Internet Penetration (%)	765	46.98	28.41	0.06	99.65
GDP Per Capita (Tho. US dollars)	782	20.56	19.17	0.44	88.42
GDP (Tri. US dollars)	782	11.29	24.60	0.05	205.44
Population (Mil. People)	801	83.94	189.41	0.59	1,311.05

Notes: This sample includes 42 countries over 2000-2018, with a total of 728 observations.

thus use web traffic data to trace out the location of AliExpress buyers and use this geographical heterogeneity to test the impact of AliExpress. In principle, if AliExpress is large enough, we would expect Chinese exports in the apparel sector to flow towards those countries with larger shares of AliExpress visitors. Figure 3 reports the map of the number of visitors on AliExpress.com by country expressed as a percentage of total worldwide visitors. The countries with no data are the countries which, given the low number of AliExpress visitors, Alexa does not report. Continuing, you can see the difference in the share of AliExpress visitors between Argentina and Mexico. You can also see how the United States and Russia are the two countries with larger values. For our analysis, we will assume that the ranking of countries by this variable did not change (or did not change much) over 2010-2018. In addition, we shall use the natural logarithm of the share of AliExpress visitors in our regressions. The density function and histogram of this variable is reported by Figure A2 in the Appendix.

The share of AliExpress visitors may correlate with other macroeconomic indicators, such as country's GDP and level of internet penetration, which in turn can affect trade flows. Table 2 reports a correlation table with the natural logarithm of the share of AliExpress visitors and a set of candidates for alternate explanations and trade controls. As you can see, our variable of interest is positively correlated with GDP (column 2), population (column 4), whether a country shares a border with China (column 8), and possibly internet penetration (column 10), as expected. Against our expectations, the share of AliExpress visitors is negatively correlated

Figure 3: Visitors on AliExpress.com



Notes: Figure 3 reports the map of the number of visitors on AliExpress.com by country expressed as a percentage of total worldwide visitors. Observations are divided in 6 classes: the first one includes countries with no data, the following four are computed with quartiles on a dataset which excludes the two maximum values (United States and Russia), and the last one contains the two countries with maximum values (United States and Russia).

with whether a country has a free trade agreement with China (column 5) and whether it has an official language in common with China (column 6).

Table 2: Correlation

	Dependent variable:									
	ln(AliExpress$_c$)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln(Internet$_c$)	0.292									0.749*
	(0.216)									(0.371)
ln(GDP$_c$)		0.427***								0.422
		(0.063)								(1.489)
ln(GDPCap$_c$)			0.143							−0.256
			(0.106)							(1.545)
ln(POP$_c$)				0.375***						0.011
				(0.082)						(1.430)
ln(Distance$_c$)					−0.227					−0.064
					(0.195)					(0.196)
FTA$_c$						−0.668***				−0.349*
						(0.188)				(0.183)
Language$_c$							−1.018***			−0.311
							(0.127)			(0.301)
Border$_c$								1.040*		0.992**
								(0.596)		(0.459)
Legal$_c$									0.148	0.056
									(0.278)	(0.200)
Constant	−5.516***	−4.970***	−4.676***	−5.624***	−3.844***	−4.225***	−4.280***	−4.379***	−4.332***	−7.339
	(0.876)	(0.116)	(0.279)	(0.251)	(0.392)	(0.137)	(0.127)	(0.120)	(0.145)	(6.082)
SE Cluster	c	c	c	c	c	c	c	c	c	c
Observations	42	42	42	41	42	42	42	42	42	41
Adjusted R^2	0.004	0.520	0.015	0.338	−0.001	0.050	0.014	0.090	−0.020	0.611

Notes: Table 2 reports correlation coefficients between ln(AliExpress) and predictiors of a country's adoption of AliExpress. "Internet" is internet penetration, "ln(GDP)" is the natural logarithm of a country's GDP, "ln(GDPcap)" is the natural logarithm of a country's GDP per capita, "ln(POP)" is the natural logarithm of a country's population,"ln(Distance)" is the natural logarithm of a country's distance from China. "FTA", "Language", "Border" and "Legal" are dummy variables which equal one if a country has, respectvely, either a free trade agreement, a common official language, a common border or a common legal system with China. Standard errors culstered by countries are reported in parenthesis. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

3.3 Multiple Sectors

In Section 5 we extend the analysis to multiple sectors. Export data on all sectors are always from Comtrade, with around 97 HS-2 categories of goods. Summary statistics by sector are reported in Table A1 in the Appendix. For this analysis, we define goods that can be sold on AliExpress vis-à-vis goods that cannot. To do so, we consider the classification of goods proposed by AliExpress.com today and we build a dummy variable which equals 1 if we can find a correspondence with an HS-2 category and 0 otherwise. Table A1 in the Appendix shows the thus-defined classification of HS-2 categories. As few categories are not univocally traded on AliExpress, in the robustness checks we use different versions of this bivariate variable.[3]

4 Baseline Results in the Apparel Sector

We formally test this hypothesis with a generalised difference-in-differences (DID) analysis on the panel of 42 trade partners, for which web traffic data for AliExpress.com is available. Specifically, we regress Chinese gross exports in the apparel sector on an interaction term between a time dummy which equals one after the introduction of AliExpress in 2010 and a continuous "treatment" variable measuring the number of AliExpress visitors by trade partner expressed as a percentage of worldwide visitors on AliExpress. We thus estimate the following generalised difference-in-differences (DID) model for the apparel sector:

$$\begin{aligned} Export_{ct} = {} & \beta(Post2010_t \times ln(AliExpress_c)) + Post2010_t \times \boldsymbol{X}'_c\boldsymbol{\gamma} + \boldsymbol{X}_{ct}\boldsymbol{\lambda} \\ & + \boldsymbol{\delta}_t + \boldsymbol{\delta}_c + \alpha + \epsilon_{ct} \end{aligned} \tag{1}$$

where "Export" is Chinese export in the apparel sector (HS-code 61) for trade partner "c" in year "t", "Post2010" is a time dummy which equals 1 after 2010 (year of introduction of AliExpress), "AliExpress" is the share of AliExpress visitors in trade partner "c", $\boldsymbol{X'}$ is a set of time invariant controls for trade partner "c", $\boldsymbol{X}$ is a set of time varying controls for trade partner "c" in year "t", $\boldsymbol{\delta}$ are year and country fixed effects, α is a constant, and ϵ is the error term. We expect a positive and significant β, as the post-2010 increases in Chinese exports in the apparel sector should be larger for those trading partners with larger shares of AliExpress visitors.

Table 3 reports the coefficient estimates. Column 1 reports the simple OLS estimate, column 2 includes country and time fixed effects, column 3 adds controls, column 4 shows results when the US is excluded from our sample, and column 5 reports the results when we use total share of Chinese exports as the dependent variable. As you can see, the coefficient of the interaction term of interest is positive and statistically significant in the reported specifications. If we interpret the coefficient in column 3, we can say that, when moving from a trade partner at the 10th to 90th percentile of the distribution of the share of AliExpress visitors, the post-2010 increase in Chinese

[3] See Table A3 in the Appendix.

exports in the apparel sector is about 32 million \$ larger $(1.01 \times (.037 - .006) \times 1000)$. When we drop the United States, which has a large share of AliExpress visitors and is also a major destination of Chinese exports, this difference (in differences) goes down to about 15 million \$ $(0.493 \times (.035 - .006) \times 1000)$. We pick the more conservative specification of column (4) as our preferred specification. The robustness checks for this specification are reported in Table A2 in the Appendix.

Table 3: Baseline

	Dependent variable:				
	Export$_{ct}$				Share$_{ct}$
	(1)	(2)	(3)	(4)	(5)
Post2010$_t$ × ln(AliExpress$_c$)	0.978**	0.942**	1.010**	0.493***	0.060***
	(0.439)	(0.450)	(0.428)	(0.121)	(0.015)
Post2010$_t$ × ln(Distance$_c$)			0.140	−0.009	−0.001
			(0.273)	(0.216)	(0.026)
Post2010$_t$ × FTA$_{ct}$			1.014**	0.617***	0.075***
			(0.432)	(0.204)	(0.025)
Post2010$_t$ × Language$_c$			−0.323	−0.345	−0.042
			(0.326)	(0.253)	(0.031)
Post2010$_t$ × Border$_c$			−1.000*	−0.431*	−0.052*
			(0.538)	(0.250)	(0.030)
Post2010$_t$ × Legal$_c$			−0.113	0.146	0.018
			(0.338)	(0.240)	(0.029)
ln(Internet$_{ct}$)			−0.369**	−0.220***	−0.027***
			(0.167)	(0.066)	(0.008)
ln(GDP$_{ct}$)			1.800	2.208	0.268
			(2.259)	(1.678)	(0.204)
ln(POP$_{ct}$)			−2.149	−2.773	−0.336
			(2.572)	(1.958)	(0.238)
ln(GDPCap$_{ct}$)			−2.420	−2.635	−0.320
			(2.334)	(1.840)	(0.223)
FTA$_{ct}$			−0.697**	−0.435***	−0.053***
			(0.303)	(0.144)	(0.018)
ln(AliExpress$_c$)	1.055**				
	(0.494)				
Post2010$_t$	4.811**				
	(2.027)				
Constant	5.256**				
	(2.320)				
FE	-	c,t	c,t	c,t	c,t
SE Cluster	c	c	c	c	c
N Countries	42	42	42	41	41
N Years	20	20	20	20	20
Observations	812	812	756	738	738
Adjusted R^2	0.294	0.885	0.897	0.933	0.933
Note:					*p<0.1; **p<0.05; ***p<0.01

Notes: Table 3 reports baseline specifications for the link between the use of AliExpress.com by a trade partner and Chinese exports. The reference sample is based on data of Chinese exports in the apparel sector (HS-2 code 61) for 42 trade partners, over 2000-2018. Standard errors clustered by countries are reported in parenthesis. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

5 Extensions

5.1 Generalization to All Sectors

In this section we expand our baseline results to all sectors of Chinese exports. To do so, we hand-code a sector-level dummy which equals 1 if we can find a match of the respective HS-2 classification on the current categories of goods sold of AliExpress.com, and 0 otherwise.[4] We thus first estimate our model in Equation 4 for sectors that can be traded on AliExpress and sectors that cannot, as we expect the coefficient of the first sample to be larger. Secondly, we test whether this difference is statistically significant by interacting our DID term with the sector-level dummy, as follows:

$$Export_{cts} = \theta(Post2010_t \times ln(AliExpress_c) \times D_s) + Post2010_t \times \boldsymbol{X}_c'\boldsymbol{\gamma} + \boldsymbol{X}_{ct}\boldsymbol{\lambda} \\ + \boldsymbol{\delta}_t + \boldsymbol{\delta}_c + \boldsymbol{\delta}_s + \alpha + \epsilon_{cts} \tag{2}$$

where D_s is the sector-level dummy. A positive and significant θ coefficient will indicate that the mentioned difference is statistically significant at conventional standards.

Table 4 reports the model estimates for more sectors of Chinese exports. Column 1 reports the baseline specification for the apparel sector (column 4 of Table 3). Then, as you can see from columns 2 and 3, we find that the DID coefficient estimate is larger for sectors that can be traded on AliExpress than for sectors which cannot be traded on AliExpress. Two things are worth noting here. First, the coefficient for the apparel sector (.493 in column 1) is more than twice the coefficient for all AliExpress sectors (.227 in column 2). This is expected, as the AliExpress impact is neater in the apparel sector, the leading sector on AliExpress. Second, column 3 shows that countries with larger shares of AliExpress visitors also imported more non-AliExpress goods after 2010. This being given, in column 4 we estimate Equation 2 and we test whether this difference in DID terms between AliExpress and non-AliExpress sectors is also statistically significant. As you can see from the triple DID coefficient estimate of .043 in column 4, it is. Here we have reported the specification with country, year and sector fixed effects. In Table

[4] Table A1 in the Appendix shows which HS-2 codes can be traded on AliExpress by this definition. As a precise matching is not always possible, in Table A3 of the Appendix we test whether our results hold with two other versions of this variable.

A3 we also show that results hold when we include country-year and sector fixed effects, which reduce issues related to omitted variable bias and reverse causality at the country-year level.

Table 4: Sector Comparison

	Dependent variable:			
	Export			
	Apparel	Ali Sectors	No Ali Sectors	Ali vs No Ali
	(1)	(2)	(3)	(4)
$\text{Post2010}_t \times \ln(\text{AliExpress}_c)$	0.493***	0.227***	0.076***	0.043***
	(0.121)	(0.061)	(0.023)	(0.014)
$\text{Post2010}_t \times \ln(\text{AliExpress}_c) \times \text{D}_s$				0.073***
				(0.017)
$\text{Post2010}_t \times \text{D}_s$				0.408***
				(0.080)
$\ln(\text{AliExpress}_c)$ times D_s				0.103**
				(0.041)
$\text{Post2010}_t \times \ln(\text{Distance}_c)$	−0.009	−0.134	−0.033	−0.016
	(0.216)	(0.084)	(0.031)	(0.018)
$\text{Post2010}_t \times \text{FTA}_{ct}$	0.617***	0.134	0.011	0.037
	(0.204)	(0.122)	(0.049)	(0.030)
$\text{Post2010}_t \times \text{Language}_c$	−0.345	0.129	0.106**	0.018
	(0.253)	(0.121)	(0.051)	(0.033)
$\text{Post2010}_t \times \text{Border}_c$	−0.431*	−0.050	0.074	0.034
	(0.250)	(0.165)	(0.079)	(0.054)
$\text{Post2010}_t \times \text{Legal}_c$	0.146	0.113	0.024	0.015
	(0.240)	(0.108)	(0.039)	(0.022)
$\ln(\text{Internet}_{ct})$	−0.220***	−0.098***	−0.029**	−0.035***
	(0.066)	(0.031)	(0.012)	(0.009)
$\ln(\text{GDP}_{ct})$	2.208	0.212	−0.069	0.055
	(1.678)	(0.618)	(0.224)	(0.167)
$\ln(\text{POP}_{ct})$	−2.773	−0.346	0.049	−0.065
	(1.958)	(0.711)	(0.260)	(0.201)
$\ln(\text{GDPCap}_{ct})$	−2.635	−0.331	0.036	−0.052
	(1.840)	(0.668)	(0.237)	(0.171)
FTA_{ct}	−0.435***	−0.111*	−0.028	−0.043**
	(0.144)	(0.067)	(0.027)	(0.018)
FE	c,t	c,t	c,t	c,t,s
SE Cluster	c	c	c	c
N Countries	41	41	41	41
N Years	20	20	20	20
N Sectors	1	32 (A)	65 (A)	97
Observations	738	738	738	69,068
Adjusted R^2	0.933	0.920	0.901	0.490

Notes: Table 4 reports a generalisation of our findings to more sectors. "D" is a sector dummy which equals one if a good "s" can be sold on AliExpress and zero otherwise. "Post2010" is a time dummy which equals one after 2010, the year of introduction of AliExpress. The reference sample is constructed with data on Chinese exports at the 2-HS level for 42 trade partners over 2000-2019. Standard errors culstered by countries are reported in parenthesis. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Given the results for the apparel sector and multiple sectors, in the Appendix we test whether the AliExpress impact is large enough to "defeat gravity" by estimating a gravity equation. We do so with two important caveats. First, in a gravity analysis we need to assume that no other countries has a platform like AliExpress, i.e. all other countries are controls. Second, we do not dispose of AliExpress data and therefore we cannot perform the exercise of comparing AliExpress flows with aggregate trade data, as it is usually done in the literature (e.g. Lendle et al., 2016; Chen and Wu, 2020). Given these two caveats, we do not find compelling evidence that the AliExpress impact is large enough to defeat gravity, as it is shown in Table A4 in the Appendix.[5]

5.2 Exogeneity

In this section we address the possible sources of endogeneity. First, we can have omitted variable bias, as, for example, some countries may have experienced a post-2010 boost in parcel infrastructure, which makes it easier for people to get goods shipped from China. This could be in turn an incentive for consumers from those countries to create an account on AliExpress. Second, some countries may just import more apparel goods from China after 2010, following the increase in, say, number of Chinese shops and business in these countries. We may have an overestimation if consumers in these countries start to use AliExpress as they find it easier to order online rather than going to the local shop. On the other hand, we may underestimate the AliExpress effect if consumers in these countries still prefer to go in person to the local shop to buy Chinese goods rather than order them online on AliExpress.

To address these issues, we use an instrumental-variable (IV) identification strategy which exploits the cross-country spillovers in the use of AliExpress. The literature shows that some spillovers are possible. For example, Blum and Goldfarb (2006) show that Americans are more likely to visit websites that are hosted in nearby countries, even after controlling for language, income, and immigrant shock. In our

[5] Table A4 is an extension to a gravity set up. We confirm a negative coefficient of distance in our sample (negative coefficient in column 1). We also find that distance plays a smaller role for China with respect to the other 42 countries (positive coefficient in column 2, could be due to better shipping systems or other reasons). From this starting point, while we find that this is even more true in those sectors that can be traded on AliExpress (positive coefficient in column 3), this is not the case for those countries with larger shares of AliExpress (insignificant coefficients in columns 4 and 5). Bottom line: we do not find compelling evidence that AliExpress can overcome distance.

case, we focus on the spillovers that can come from the world of mouth (WOM). Roy et al. (2014) report that WOM is an important determinant of website proliferation for e-commerce platforms. In principle, via cross-country WOM, users in a country may be more likely to join AliExpress if consumers in bordering countries are already using it. To build our instrument, we follow the methodology proposed by Caselli and Reynaud (2020), who exploit the fact that a government may be induced by bordering countries to apply a fiscal reform. Our IV thus measures the number of bordering countries using AliExpress in the following way:

$$AliContig_c = \sum_{j \neq c}^{n-c} AliD_j \times ContigD_{jc}$$

where *AliD* is a dummy which equals one if country j uses AliExpress and zero otherwise, and *ContigD* is a dummy which equals one if country j shares a border with country c. We use it to run the following first and second stages:

$$\begin{aligned} Post2010_t \times AliD_c = {} & \theta Post2010 \times AliContig_c + Post2010_t \times \boldsymbol{X}_c' \boldsymbol{\gamma}_1 + \boldsymbol{X}_{ct} \boldsymbol{\lambda}_1 \\ & + \boldsymbol{\delta}_{1t} + \boldsymbol{\delta}_{1c} + \alpha_1 + \epsilon_{1ct} \end{aligned} \tag{3}$$

$$\begin{aligned} Export_{ct} = {} & \beta \widehat{Post2010_t \times} AliD_c + Post2010_t \times \boldsymbol{X}_c' \boldsymbol{\gamma}_2 + \boldsymbol{X}_{ct} \boldsymbol{\lambda_2} \\ & + \boldsymbol{\delta}_{2t} + \boldsymbol{\delta}_{2t} + \alpha_2 + \epsilon_{2ct} \end{aligned} \tag{4}$$

where *AliD* is a dummy which equals one if country c uses AliExpress and zero otherwise. Two things are worth noting here. First, in order to have a strong IV, here we use a bivariate version of our variable of interest (rather than continuous). As a result, our sample of countries increases to 228. Second, in our IV specifications we use region-time fixed effects. We do so because there may be some regional dynamics affecting the exogeneity of our instrument. For example, regions that generally import more from China may use AliExpress more. The region-time fixed effects control for these issues.

Table 5 reports the estimation results with the IV. Column 1 shows the model estimates with OLS and fixed effects. It mirrors results in 3, with the two main differences mentioned above. As you can see, we confirm our results also in this setting. Column 2 is the first stage. As we have one instrument for one endogenous variable, the F test comparing models with and without the IV is the test for relevance of the IV. Note that the value of the F statistics is 19.66, which is above the rule-of-thumb value of 10 proposed in Staiger and Stock (1997). Column 3 shows the reduced form. The positive and significant coefficient is reassuring, as it shows that some of the impact of AliExpress on Chinese exports is explained by the IV.

Table 5: Instrumental Variable

	Dependent variable:			
	OLS	1st Stage	Red. Form	2nd Stage
	Export	Post2010 × AliD	Export	Export
	(1)	(2)	(3)	(4)
$\text{Post2010}_t \times \text{AliD}_c$	0.330***			0.865**
	(0.071)			(0.357)
$\text{Post2010}_t \times \text{AliContig}_c$		0.146***	0.127**	
		(0.033)	(0.052)	
$\text{Post2010}_t \times \ln(\text{Distance}_c)$	0.052	−0.160	−0.110	0.029
	(0.087)	(0.195)	(0.103)	(0.136)
$\text{Post2010}_t \times \text{FTA}_{ct}$	0.088	−0.029	0.100	0.125
	(0.065)	(0.130)	(0.062)	(0.127)
$\text{Post2010}_t \times \text{Language}_c$	−0.642	−0.090	−0.686	−0.608
	(0.431)	(0.294)	(0.448)	(0.450)
$\text{Post2010}_t \times \text{Border}_c$	−0.013	−0.249	−0.227	−0.012
	(0.171)	(0.195)	(0.175)	(0.200)
$\text{Post2010}_t \times \text{Legal}_c$	0.061	−0.183	−0.104	0.054
	(0.147)	(0.151)	(0.130)	(0.158)
$\ln(\text{Internet}_{ct})$	−0.059***	−0.060***	−0.084***	−0.032
	(0.021)	(0.021)	(0.025)	(0.029)
$\ln(\text{GDP}_{ct})$	−0.043	0.049	0.056	0.013
	(0.345)	(0.624)	(0.309)	(0.598)
$\ln(\text{POP}_{ct})$	0.144	−0.127	−0.002	0.109
	(0.304)	(0.564)	(0.277)	(0.509)
$\ln(\text{GDPCap}_{ct})$	−0.075	−0.075	−0.213	−0.149
	(0.332)	(0.624)	(0.301)	(0.593)
FTA_{ct}	−0.033	0.052	−0.020	−0.065
	(0.077)	(0.040)	(0.072)	(0.093)
1st Stage F Test (Relevance)	-	19.66	-	-
FE	c,rt	c,rt	c,rt	c,rt
SE Cluster	c	c	c	c
N Countries	228	228	228	228
N Years	20	20	20	20
Observations	2,796	2,796	2,796	2,796
Adjusted R^2	0.942	0.587	0.941	-

Notes: Table 5 reports the instrumental-variable results. The variable of interest is "AliD", which equals 1 if a country uses AliExpress and 0 otherwise. The IV is "AliContig", which reports the number of bordering countries using AliExpress. The reference sample is constructed with data Chinese exports in the apparel sector (61 HS-code) of 228 countries over 2000-2019. We include region-time FE to reduce reverse causality due to regional dynamics. Standard errors culstered by countries are reported in parenthesis. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Finally, column 4 shows the second stage (or IV estimation of the structural model).

The positive and statistically significant coefficient shows that results hold when we use this IV, with somewhat larger effect. Overall, the IV estimation indicates that the sources of endogeneity may lead to underestimation, so the baseline results in Table 3 can be considered a lower bound.

6 AliExpress and Global Value Chains

In this second part of the paper we assess whether the AliExpress-induced increase in Chinese exports was beneficial for Chinese (rather than foreign) value added. We start by checking the degree of product differentiation of AliExpress-induced exports. To measure the level of product differentiation in a sector, we rely on the measure proposed by Rauch (1999). By Rauch's definition, a product is differentiated if is not listed on either a stock exchange or in trade books. Note that, as Rauch uses SITC-4 (version 2) classification, we had to convert SITC 4 to our HS classification. Specifically, after converting SITC 4 to HS 4 (almost 1-on-1 mapping), for each HS-2 code we have computed the percentage of HS-4 components that are classified as differentiated sectors. We use the thus obtained measure for sector differentiation to estimate the following model:

$$\begin{aligned} Export_{cts} = {} & \theta(Post2010_t \times ln(AliExpress_c) \times D_s \times Diff_s) + Post2010_t \times \boldsymbol{X}_c'\boldsymbol{\gamma} \\ & + \boldsymbol{X}_{ct}\boldsymbol{\lambda} + \boldsymbol{\delta}_t + \boldsymbol{\delta}_c + \boldsymbol{\delta}_s + \alpha + \epsilon_{cts} \end{aligned} \tag{5}$$

A positive and significant estimate of the θ coefficient would indicate that the AliExpress-induced increase in Chinese exports is larger for differentiated products, which should contain a large share of Chinese value added.

Table 6 report the coefficient estimates of Equation 5.[6] Column 1 shows again the baseline results with multiple sectors, with coefficient estimate .073 (Table 4, column 4). The positive and significant coefficient of .001 in column 2 shows that the triple-interaction term of interest is larger for highly differentiated sectors. In column 3 we also show that statistical significance at statistical standards goes away when we use another version of this index, based on Rauch's liberal classification (rather than conservative). Overall, Table 6 brings some evidence that AliExpress supported differentiated sectors and therefore Chinese factors and value added.

As a further test for this hypothesis, we re-estimate the specifications in Table 4

[6] For presentation purposes, only coefficients of interest are reported

Table 6: Differentiation Index (Rauch, 1999)

	Dependent variable:		
	Export		
	(1)	(2)	(3)
$\text{Post2010}_t \times \ln(\text{AliExpress}_c) \times \text{D}_s$	0.073^{***}	-0.052^{*}	-0.0003
	(0.017)	(0.029)	(0.034)
$\text{Post2010}_t \times \ln(\text{AliExpress}_c) \times \text{D}_s \times \text{Diff1}_s$		0.001^{**}	
		(0.0004)	
$\text{Post2010}_t \times \ln(\text{AliExpress}_c) \times \text{D}_s \times \text{Diff2}_s$			0.0004
			(0.0005)
FE	c,t,s	c,t,s	c,t,s
SE Cluster	c	c	c
N Countries	41	41	41
N Years	20	20	20
N Sectors	97	97	97
Observations	69,068	69,068	69,068
Adjusted R^2	0.490	0.496	0.495

Notes: Table 6 reports the country-time-sector level analysis, extended to include an index measuring the level of product differentiation by HS-2 sectors. "Diff1" is the percentage of differentiated goods contained in each HS-2 sector, following the conservative definition in Rauch (1999). "Diff2" is an alternative version of the same percentage, computed follwing the liberal definition in Rauch (1999). Standard errors culstered by countries are reported in parenthesis. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

using export data from the dataset Trade in Value Added (TiVA) by the The Organisation for Economic Co-operation and Development (OECD). Among others, TiVA includes Chinese exports in domestic value added over 2005-2015 ("EXGR_DVA"). The definitions of export sectors in Tiva are different from our starting-point HS classification. In our case, we have a total of 36 TiVA sectors, for which we have recoded our dummy measuring if a sector can be traded on AliExpress or not.
As a consequence, when using TiVA we only have 4 sectors that can be traded on AliExpress, and 32 that cannot. Table 7 reports the estimation results. Columns (1)-(4) reproduce the results of Table 4 using Chinese gross exports from TiVA. The columns' order is the same of Table 4 (apparel, sectors that can be traded on AliExpress, sectors that cannot be traded on AliExpress, and all sectors). As you can see, the positive and significant estimates confirm our former results. One thing to note is that estimates are somewhat larger than in Table 4. This is because

Table 7: AliExpress and Domestic Value Added

	Dependent variable:							
	Export				Export in DVA			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{Post2010}_t \times \ln(\text{AliExpress}_c)$	1.722***	1.305***	0.165**	0.140**	1.559***	1.086***	0.142**	0.122**
	(0.484)	(0.460)	(0.065)	(0.062)	(0.446)	(0.369)	(0.055)	(0.052)
$\text{Post2010}_t \times \ln(\text{AliExpress}_c) \times \text{D}_s$				0.804***				0.708***
				(0.163)				(0.129)
$\text{Post2010}_t \times \ln(\text{Distance}_c)$	−1.366	−1.418**	−0.272***	−0.251***	−1.279	−1.149**	−0.232***	−0.222***
	(0.944)	(0.694)	(0.096)	(0.085)	(0.881)	(0.566)	(0.084)	(0.075)
$\text{Post2010}_t \times \text{FTA}_{ct}$	1.361*	1.083*	0.182*	0.169**	1.288*	0.920*	0.156*	0.155**
	(0.806)	(0.625)	(0.093)	(0.080)	(0.762)	(0.523)	(0.081)	(0.070)
$\text{Post2010}_t \times \text{Language}_c$	−2.886**	−2.862***	−0.393***	−0.543***	−2.691**	−2.305***	−0.339***	−0.469***
	(1.385)	(1.024)	(0.136)	(0.174)	(1.277)	(0.828)	(0.117)	(0.152)
$\text{Post2010}_t \times \text{Border}_c$	−1.610	−1.355	−0.045	0.085	−1.437	−1.123	−0.046	0.040
	(1.606)	(1.423)	(0.227)	(0.246)	(1.473)	(1.136)	(0.193)	(0.202)
$\text{Post2010}_t \times \text{Legal}_c$	−0.150	0.450	−0.014	−0.012	−0.156	0.298	−0.012	−0.011
	(0.825)	(0.763)	(0.092)	(0.116)	(0.753)	(0.609)	(0.079)	(0.102)
$\ln(\text{Internet}_{ct})$	−1.096	−0.198	−0.020	−0.138	−1.061	−0.264	−0.023	−0.121
	(1.029)	(1.066)	(0.116)	(0.164)	(0.941)	(0.846)	(0.101)	(0.141)
$\ln(\text{GDP}_{ct})$	−9.064	−5.921	−2.548**	−1.793	−8.032	−5.385	−2.174**	−1.654
	(10.427)	(11.501)	(1.266)	(1.594)	(9.355)	(8.865)	(1.084)	(1.357)
$\ln(\text{POP}_{ct})$	6.662	8.460	3.229**	2.887	5.767	7.082	2.721**	2.547
	(10.088)	(11.739)	(1.376)	(1.892)	(9.016)	(9.149)	(1.184)	(1.617)
$\ln(\text{GDPCap}_{ct})$	11.564	7.420	2.746**	2.181	10.160	6.464	2.337**	1.950
	(10.619)	(11.530)	(1.249)	(1.593)	(9.522)	(8.864)	(1.070)	(1.357)
FTA_{ct}	−0.977**	−0.906***	−0.146***	−0.210***	−0.918***	−0.755***	−0.127***	−0.183***
	(0.383)	(0.329)	(0.038)	(0.054)	(0.349)	(0.255)	(0.033)	(0.046)
$\text{Post2010}_t \times \text{D}_s$				4.653***				4.055***
				(0.729)				(0.583)
$\ln(\text{AliExpress}_c) \times \text{D}_s$				1.600***				1.133***
				(0.418)				(0.298)
Sector	Apparel	Ali	No Ali	Ali vs No Ali	Apparel	Ali	No Ali	Ali vs No Ali
FE	c,t	c,t	c,t	c,t,s	c,t	c,t	c,t	c,t,s
SE Cluster	c	c	c	c	c	c	c	c
N Countries	31	31	31	31	31	31	31	31
N Years	11	11	11	11	11	11	11	11
N Sectors	1	4 (A)	32 (A)	36	1	4 (A)	32 (A)	36
Observations	341	341	341	12,276	341	341	341	12,276
Adjusted R^2	0.957	0.936	0.938	0.586	0.953	0.930	0.936	0.576

Notes: Table 7 extends the analysis to GVC variables. "Export" is gross Chinese exports from TiVA (EXGR). "Export in DVA" is gross Chinese exports expressed in terms of domestic value added from TiVA (EXGR_DVA). The sample includes Chinese exports to 31 of the 42 countries for which we have AliExpress visitors, over 2005-2015. Standard errors culstered by countries are reported in parenthesis. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

TiVA's sectors include more goods than our HS classification. For example, the average China export in our sample in sector HS 61 (apparel) is around 1 billion \$, while the average China export in the correspondent TiVA sector (textile) is 5 billion \$. Columns (5)-(8) show that the results hold when we consider China exports expressed in terms of domestic value added. It was largely expected for textile (column 5), as around 86% of the value added in textile is Chinese. Columns 6 and 7 shows that results hold also when we compare sectors that can be traded on AliExpress and sectors that cannot, as the coefficient of columns 6 is larger than the coefficient of column 7. Column 8 shows that this difference is also statistically significant (triple interaction term). Overall, Table 7 brings some evidence in favour of the idea that the AliExpress helped export with a significant share of Chinese value added.

7 Conclusions

We make three main contributions in this paper. The first is to provide a line-sketch of a theory model that emphasizes the value-creating aspect of superior matching between consumers preferences and the varieties when they can purchase online (and thus have access to a broader range of varieties than is available locally). Second, to the best of our knowledge, we are the first to empirically estimate the impact that an online-trade platform can have on the gross export of an emerging economy. Third, we provide evidence that the introduction of such platforms can support export of domestic value added.

Our baseline results show that the post-2010 increase in Chinese exports in the apparel sector is larger for trade partners with larger shares of AliExpress visitors. These results are robust to different checks and are confirmed also when we extend the analysis to more sectors of Chinese exports. We also find that the AliExpress-induced boost in export came to the advantage of Chinese factors. We argue that export-oriented platforms like AliExpress can play an important role in the development path of emerging economies and help them moving up in the global value chain.

Given our data limitations, our findings are only supportive of the mechanisms we highlight. In future work, we shall strive to get more firm-level data on AliExpress so as to be able to more precisely investigate the role of matching in the trade creation observed.

References

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

James E Anderson and Douglas Marcouiller. Insecurity and the pattern of trade: An empirical investigation. *Review of Economics and statistics*, 84(2):342–352, 2002.

Jie Bai, Maggie Xiaoyang Chen, and Daniel Xu. Search and information frictions on global e-commerce platforms: Evidence from aliexpress. 2018. NET Institute Working Paper No. 18-17.

Richard Baldwin and James Harrigan. Zeros, quality, and space: Trade theory and trade evidence. *American Economic Journal: Microeconomics*, 3(2):60–88, 2011.

Richard Baldwin and Anthony J Venables. Spiders and snakes: Offshoring and agglomeration in the global economy. *Journal of International Economics*, 90(2):245–254, 2013.

Mina N Baliamoune. The new economy and developing countries: Assessing the role of ict diffusion. 2002. WIDER Discussion Paper, No. 2002/77.

Bernardo S Blum and Avi Goldfarb. Does the internet defy the law of gravity? *Journal of international economics*, 70(2):384–405, 2006.

Jeffrey R Brown and Austan Goolsbee. Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of political economy*, 110(3):481–507, 2002.

Erik Brynjolfsson and Michael D Smith. Frictionless commerce? a comparison of internet and conventional retailers. *Management science*, 46(4):563–585, 2000.

Erik Brynjolfsson, Xiang Hui, and Meng Liu. Does machine translation affect international trade? evidence from a large digital platform. *Management Science*, 65(12):5449–5460, 2019.

Jerónimo Carballo, Alejandro Graziano, Georg Schaur, and Christian Volpe Martincus. The border labyrinth: Information technologies and trade in the presence of multiple agencies. 2016. IDB Working Paper Series, No. IDB-WP-706.

Francesca Caselli and Julien Reynaud. Do fiscal rules cause better fiscal balances? a new instrumental variable strategy. *European Journal of Political Economy*, page 101873, 2020.

Thomas Chaney. The network structure of international trade. *American Economic Review*, 104(11):3600–3634, 2014.

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

George RG Clarke and Scott J Wallsten. Has the internet increased trade? developed and developing country evidence. *Economic Inquiry*, 44(3):465–484, 2006.

Victor Couture, Benjamin Faber, Yizhen Gu, and Lizhi Liu. E-commerce integration and economic development: Evidence from china. 2018. NBER Working Paper 24384.

Kunal Dasgupta and Jordi Mondria. Inattentive importers. *Journal of International Economics*, 112:150–165, 2018.

Haluk Demirkan, Michael Goul, Robert J Kauffman, and David M Weber. Does distance matter? the influence of ict on bilateral trade flows. 2009. GlobDev 2009. 17.

Néstor Duch-Brown and Bertin Martens. Search costs, information exchange and sales concentration in the digital music industry. 2014. Institute for Prospective Technological Studies Digital Economy Working Paper No. 2014/09.

Liran Einav, Chiara Farronato, and Jonathan Levin. Peer-to-peer markets. *Annual Review of Economics*, 8:615–635, 2016.

Jingting Fan, Lixin Tang, Weiming Zhu, and Ben Zou. The alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce. *Journal of International Economics*, 114:203–220, 2018.

Ying Fan, Jiandong Ju, and Mo Xiao. Reputation premium and reputation management: Evidence from the largest e-commerce platform in china. *International Journal of Industrial Organization*, 46:63–76, 2016.

Chiara Farronato and Andrey Fradkin. The welfare effects of peer entry in the accommodation market: The case of airbnb. 2018. National Bureau of Economic Research, Working Paper 24361.

Ana M Fernandes, Aaditya Mattoo, Huy Nguyen, and Marc Schiffbauer. The internet and chinese exports in the pre-ali baba era. *Journal of Development Economics*, 138: 57–76, 2019.

Caroline Freund, Alen Mulabdic, and Michele Ruta. Is 3d printing a threat to global trade? the trade effects you didn't hear about. 2019. World Bank, Policy Research Working Paper N. 9024.

Caroline L Freund and Diana Weinhold. The effect of the internet on international trade. *Journal of international economics*, 62(1):171–189, 2004.

Anindya Ghose and Yuliang Yao. Using transaction prices to re-examine price dispersion in electronic markets. *Information Systems Research*, 22(2):269–288, 2011.

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

Maris Goldmanis, Ali Hortaçsu, Chad Syverson, and Önsel Emre. E-commerce and the market structure of retail industries. *The Economic Journal*, 120(545):651–682, 2010.

Gene M Grossman and Esteban Rossi-Hansberg. Trading tasks: A simple theory of offshoring. *American Economic Review*, 98(5):1978–97, 2008.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. Cultural biases in economic exchange? *The quarterly journal of economics*, 124(3):1095–1131, 2009.

Jonas Hjort and Jonas Poulsen. The arrival of fast internet and employment in africa. *American Economic Review*, 109(3):1032–79, 2019.

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

Xiang Hui. Facilitating inclusive global trade: Evidence from a field experiment. *Management Science*, 66(4):1737–1755, 2020.

Robert C Johnson. Measuring global value chains. *Annual Review of Economics*, 10: 207–236, 2018.

Chungsang Tom Lam and Meng Liu. Demand and consumer surplus in the on-demand economy: the case of ride sharing. *Social Science Electronic Publishing*, 17(8):376–388, 2017.

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

Gregory Lewis. Asymmetric information, adverse selection and online disclosure: The case of ebay motors. *American Economic Review*, 101(4):1535–46, 2011.

Fuyi Li, Stacey Frederick, and Gary Gereffi. E-commerce and industrial upgrading in the chinese apparel value chain. *Journal of Contemporary Asia*, 49(1):24–53, 2019.

Lingfang Ivy Li, Steven Tadelis, and Xiaolan Zhou. Buying reputation as a signal of quality: Evidence from an online marketplace. 2016. National Bureau of Economic Research, Working Paper No. 22584.

Ethan Lieber and Chad Syverson. Online versus offline competition. *The Oxford handbook of the digital economy*, 189, 2012.

Meng Liu, Erik Brynjolfsson, and Jason Dowlatabadi. Technology, incentives, and service quality: the case of taxis and uber. 2018. National Bureau of Economic Research, Conference Paper 113125.

Laura Márquez-Ramos and Inmaculada Martínez-Zarzoso. The effect of technological innovation on international trade. *Economics: the open-access, open-assessment E-Journal*, 4:11, 2010.

Anselm Mattes, Philipp Meinen, and Ferdinand Pavel. Goods follow bytes: The impact of ict on eu trade. 2012. DIW Berlin Discussion Paper No. 1182.

Eric Overby and Chris Forman. The effect of electronic commerce on geographic purchasing patterns and price dispersion. *Management Science*, 61(2):431–453, 2015.

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

Alberto Portugal-Perez and John S Wilson. Export performance and trade facilitation reform: Hard and soft infrastructure. *World development*, 40(7):1295–1307, 2012.

James E Rauch. Networks versus markets in international trade. *Journal of international Economics*, 48(1):7–35, 1999.

Sanjit Kumar Roy, Walfried M Lassar, and Gul T Butaney. The mediating impact of stickiness and loyalty on word-of-mouth promotion of retail websites. *European Journal of Marketing*, 48(9/10):1828–1849, 2014.

Douglas Staiger and James H Stock. Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586, 1997.

Claudia Steinwender. Information frictions and the law of one price:"when the states and the kingdom became united". 2014. WTO Staff Working Paper, No. ERSD-2014-12.

Jonathan Timmis. The internet and international trade in goods. 2012. The University of Nottingham, Discussion Paper No. 12/03.

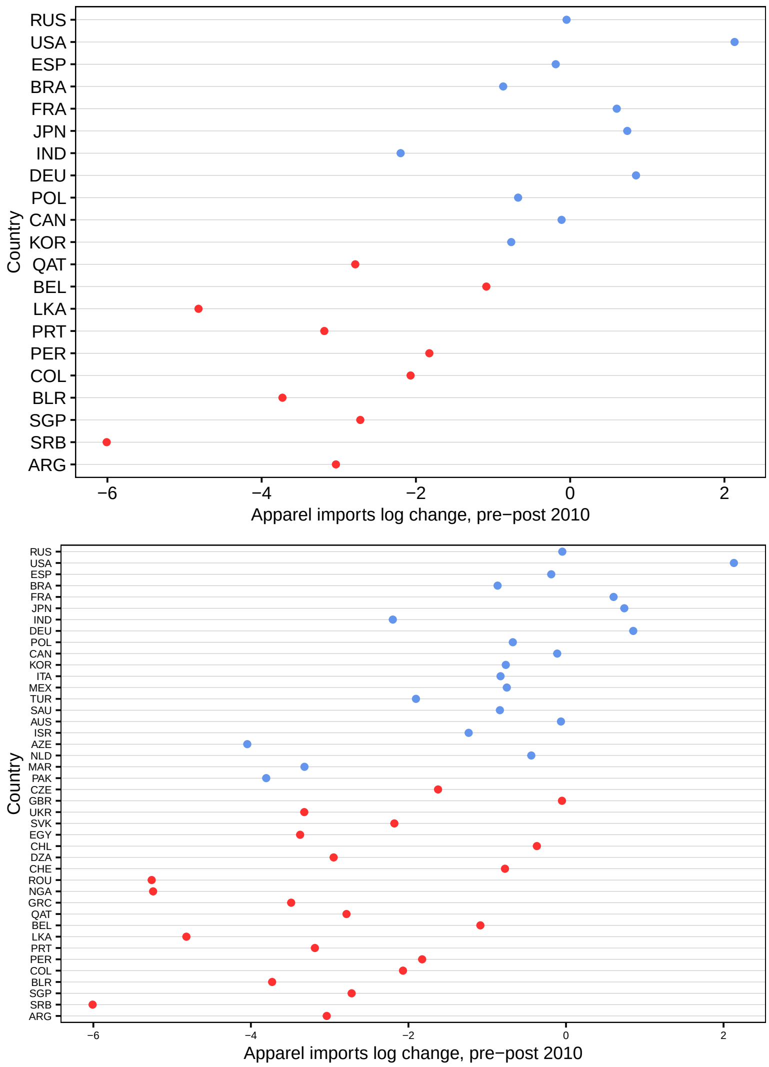
Vijay K Vemuri and Shahid Siddiqi. Impact of commercialization of the internet on international trade: A panel study using the extended gravity model. *The International Trade Journal*, 23(4):458–484, 2009.

World Bank. World development report 2020: Trading for development in the age of global value chains. 2020. The World Bank, ISBN 9781464814570.

Feng Zhu and Qihong Liu. Competing with complementors: An empirical look at amazon.com. *Strategic Management Journal*, 39(10):2618–2642, 2018.

Appendix

Figure A1: Before and after AliExpress in the Apparel Sector



Notes: Figure A1 shows the log difference in Chinese exports in the apparel sectors (x axis), with trade partners ordered by their share of AliExpress visitors. The top panel shows the trade partners below (red) and above (blue) the 25th and the 75th percentiles of the AliExpress variable. The bottom panel shows the trade partners below (red) and above (blue) the 50th percentile of the AliExpress variable (full distribution of countries). Greece and Serbia are missing as they report negative changes.

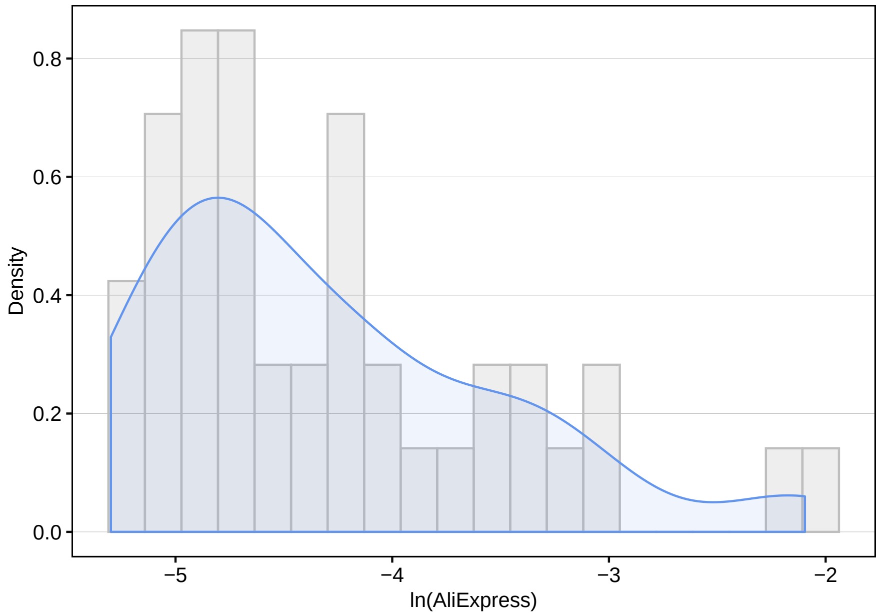
Table A1: List of HS2 sectors ordered by AliExpress Dummy

HS2 Code	Commodity Name	D	Export
01	Animals; live	0	0.075
02	Meat and edible meat offal	0	0.071
03	Fish and crustaceans, molluscs and other aquatic invertebrates	0	7.775
04	Dairy produce; birds' eggs; natural honey	0	0.301
05	Animal originated products; not elsewhere specified or included	0	1.588
06	Trees and other plants, live; bulbs, roots and the like; cut flowers and ornamental foliage	0	0.325
07	Vegetables and certain roots and tubers; edible	0	4.413
08	Fruit and nuts, edible; peel of citrus fruit or melons	0	1.552
09	Coffee, tea, mate and spices	0	1.764
10	Cereals	0	0.376
11	Products of the milling industry; malt, starches, inulin, wheat gluten	0	0.264
12	Oil seeds and oleaginous fruits; industrial or medicinal plants; straw and fodder	0	2.258
13	Lac; gums, resins and other vegetable saps and extracts	0	0.830
14	Vegetable plaiting materials; vegetable products not elsewhere specified or included	0	0.190
15	Animal or vegetable fats and oils; prepared animal fats; animal or vegetable waxes	0	0.495
16	Meat, fish or crustaceans, molluscs or other aquatic invertebrates; preparations thereof	0	4.357
17	Sugars and sugar confectionery	0	0.742
18	Cocoa and cocoa preparations	0	0.201
19	Preparations of cereals, flour, starch or milk; pastrycooks' products	0	1.108
20	Preparations of vegetables, fruit, nuts or other parts of plants	0	4.222
21	Miscellaneous edible preparations	0	1.688
22	Beverages, spirits and vinegar	0	0.482
23	Food industries, residues and wastes thereof; prepared animal fodder	0	2.035
24	Tobacco and manufactured tobacco substitutes	0	0.549
25	Salt; sulphur; earths, stone; plastering materials, lime and cement	0	2.744
26	Ores, slag and ash	0	0.456
27	Mineral fuels, mineral oils; bituminous substances; mineral waxes	0	12.836
28	Inorganic chemicals; organic and inorganic compounds of precious metals	0	13.197
29	Organic chemicals	0	47.551
30	Pharmaceutical products	0	5.411
31	Fertilizers	0	4.371
35	Albuminoidal substances; modified starches; glues; enzymes	0	1.468
36	Explosives; pyrotechnic products; matches; certain combustible preparations	0	0.835
38	Chemical products n.e.c.	0	11.304
39	Plastics and articles thereof	0	46.560
40	Rubber and articles thereof	0	15.855
41	Raw hides and skins (other than furskins) and leather	0	0.188
44	Wood and articles of wood; wood charcoal	0	11.141
45	Cork and articles of cork	0	0.070
47	Pulp of wood or other fibrous cellulosic material; recovered paper or paperboard	0	0.136
49	Printed books, newspapers, pictures; manuscripts, typescripts and plans	0	4.704
50	Silk	0	0.735
51	Wool, fine or coarse animal hair; horsehair yarn and woven fabric	0	1.214
52	Cotton	0	2.440
53	Vegetable textile fibres; paper yarn and woven fabrics of paper yarn	0	0.585
54	Man-made filaments; strip and the like of man-made textile materials	0	8.102
55	Man-made staple fibres	0	4.798
56	Wadding, felt and nonwovens, special yarns; twine, cordage, ropes and cables	0	3.129
68	Stone, plaster, cement, asbestos, mica or similar materials; articles thereof	0	7.576
70	Glass and glassware	0	11.512
72	Iron and steel	0	22.174

HS2 Code	Commodity Name	D	Export
75	Nickel and articles thereof	0	0.314
76	Aluminium and articles thereof	0	16.312
78	Lead and articles thereof	0	0.023
79	Zinc and articles thereof	0	0.393
80	Tin; articles thereof	0	0.203
81	Metals; n.e.c., cermets and articles thereof	0	3.470
84	Nuclear reactors, boilers, machinery and mechanical appliances; parts thereof	0	175.072
86	Railway, tramway locomotives, rolling-stock and parts thereof	0	3.020
87	Vehicles; other than railway or tramway rolling stock, and parts and accessories thereof	0	42.339
88	Aircraft, spacecraft and parts thereof	0	1.888
89	Ships, boats and floating structures	0	6.018
93	Arms and ammunition; parts and accessories thereof	0	0.362
97	Works of art; collectors' pieces and antiques	0	0.530
99	Commodities not specified according to kind	0	15.957
32	Tanning or dyeing extracts; tannins and their derivatives; paints, varnishes; putty ; inks	1	5.993
33	Essential oils and resinoids; perfumery, cosmetic or toilet preparations	1	4.626
34	Soap, organic surface-active agents; washing, lubricating, polishing preparations	1	2.254
37	Photographic or cinematographic goods	1	0.640
42	Articles of leather; saddlery and harness; travel goods, handbags and similar containers	1	23.930
43	Furskins and artificial fur; manufactures thereof	1	0.657
46	Manufactures of straw, esparto or other plaiting materials; basketware and wickerwork	1	0.978
48	Paper and paperboard; articles of paper pulp, of paper or paperboard	1	12.245
57	Carpets and other textile floor coverings	1	2.093
58	Fabrics; special woven fabrics, tufted textile fabrics, lace, tapestries, trimmings	1	1.834
59	Textile fabrics; impregnated, coated, covered or laminated	1	3.156
60	Fabrics; knitted or crocheted	1	4.157
61	Apparel and clothing accessories; knitted or crocheted	1	45.955
62	Apparel and clothing accessories; not knitted or crocheted	1	49.652
63	Textiles, made up articles; sets; worn clothing and worn textile articles; rags	1	20.572
64	Footwear; gaiters and the like; parts of such articles	1	33.319
65	Headgear and parts thereof	1	4.700
66	Umbrellas, sun umbrellas, walking-sticks, seat sticks, whips, riding crops	1	2.006
67	Feathers and down, prepared; and articles made of feather or of down; artificial flowers	1	3.358
69	Ceramic products	1	9.623
71	Pearls; precious stones; precious metals, and articles thereof; imitation jewellery; coin	1	7.832
73	Iron or steel articles	1	42.325
74	Copper and articles thereof	1	3.744
82	Tools, implements, cutlery, spoons and forks, of base metal; parts thereof, of base metal	1	12.568
83	Metal; miscellaneous products of base metal	1	15.090
85	Electrical machinery and equipment; sound recorders and reproducers	1	200.047
90	Optical, photographic, cinematographic, measuring, medical or surgical instruments	1	42.872
91	Clocks and watches and parts thereof	1	5.255
92	Musical instruments; parts and accessories of such articles	1	1.733
94	Furniture; bedding, mattresses, cushions; lamps; prefabricated buildings	1	49.482
95	Toys, games and sports requisites; parts and accessories thereof	1	45.088
96	Miscellaneous manufactured articles	1	11.541

Notes: Table A1 reports summary statistics for the 97 HS2-level sectors included in our sample. "Commodity Name" is the official description for each HS2 level, abbreviated for simplicity (source: Comtrade). "D" is the dummy variable which equals one if a HS2 sector can be traded on AliExpress, and zero otherwise (source: author's calculation). "Export" is the sum of Chinese exports across the 42 trade partners included in our sample, by HS2 sector.

Figure A2: Visitors on AliExpress.com



Notes: Figure A2 reports the histogram and density plot of log(AliExpress), i.e. the natural logarithm of the number of visitors on AliExpress.com by country expressed as a percentage of total worldwide visitors.

Table A2: Robustness Checks

	Dependent variable:						
	Export						ln(Export)
	Baseline	2W Clust	No LN	Allow Int.	Drop C	Drop 08	ln LHS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\text{Post2010}_t \times \ln(\text{AliExpress}_c)$	0.493***	0.493***	0.228***	0.226**	0.556***	1.029**	0.030
	(0.121)	(0.155)	(0.082)	(0.114)	(0.151)	(0.447)	(0.126)
$\text{Post2010}_t \times \ln(\text{Distance}_c)$	−0.009	−0.009		−0.022	−0.042	0.143	0.684***
	(0.216)	(0.202)		(0.241)	(0.250)	(0.276)	(0.201)
$\text{Post2010}_t \times \text{FTA}_{ct}$	0.617***	0.617***	0.675***	0.647***	0.660***	1.014**	−0.117
	(0.204)	(0.199)	(0.261)	(0.205)	(0.233)	(0.434)	(0.295)
$\text{Post2010}_t \times \text{Language}_c$	−0.345	−0.345	−0.682**	−0.408		−0.316	−0.291
	(0.253)	(0.228)	(0.299)	(0.290)	(0.000)	(0.332)	(0.338)
$\text{Post2010}_t \times \text{Border}_c$	−0.431*	−0.431*	−0.869**	−0.372	−0.582**	−1.031*	0.949**
	(0.250)	(0.257)	(0.392)	(0.313)	(0.291)	(0.569)	(0.434)
$\text{Post2010}_t \times \text{Legal}_c$	0.146	0.146	0.162	0.166	0.105	−0.119	0.488**
	(0.240)	(0.220)	(0.186)	(0.214)	(0.229)	(0.343)	(0.222)
$\ln(\text{Internet}_{ct})$	−0.220***	−0.220***		−0.162**	−0.197***	−0.370**	0.412***
	(0.066)	(0.062)		(0.068)	(0.070)	(0.167)	(0.139)
$\ln(\text{GDP}_{ct})$	2.208	2.208		0.772	2.177	1.845	3.578
	(1.678)	(1.794)		(2.883)	(1.780)	(2.252)	(3.390)
$\ln(\text{POP}_{ct})$	−2.773	−2.773		−1.463	−2.721	−2.190	−2.743
	(1.958)	(2.085)		(2.982)	(2.070)	(2.562)	(3.733)
$\ln(\text{GDPCap}_{ct})$	−2.635	−2.635		−1.156	−2.774	−2.445	−3.200
	(1.840)	(1.940)		(2.916)	(1.970)	(2.336)	(3.380)
FTA_{ct}	−0.435***	−0.435***	−0.093	−0.409***	−0.351**	−0.689**	−0.133
	(0.144)	(0.138)	(0.084)	(0.141)	(0.143)	(0.295)	(0.242)
$\text{Post2010}_t \times \text{Distance}_c$			−0.040				
			(0.029)				
Internet_{ct}			0.003				
			(0.004)				
GDP_{ct}			0.141***				
			(0.041)				
POP_{ct}			−0.009***				
			(0.002)				
GDPCap_{ct}			−0.005				
			(0.011)				
$\text{Post2010}_t \times \ln(\text{Internet}_{ct})$				0.237			
				(0.284)			
$\text{Post2010}_t \times \ln(\text{GDP}_{ct})$				2.727			
				(2.781)			
$\text{Post2010}_t \times \ln(\text{POP}_{ct})$				−2.492			
				(2.757)			
$\text{Post2010}_t \times \ln(\text{GDPCap}_{ct})$				−2.593			
				(2.744)			
FE	c,t	c,t	c,t	c,t	c,t	c,t	c,t
SE Cluster	c	c,t	c	c	c	c	c
N Countries	41	41	41	41	36	41	41
N Years	20	20	20	20	20	19	20
Observations	738	738	738	738	659	755	738
Adjusted R^2	0.933	0.933	0.948	0.935	0.936	0.895	0.935

Notes: Table A2 reports robustness checks for the baseline specification. Standard errors culstered by countries are reported in parenthesis. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A3: Robustness Checks with More Sectors

	Dependent variable:						
				Export			
	Baseline	Allow Int.	D Ver 2	D Ver 3	FE 1	FE 2	FE 3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post2010$_t$ × ln(AliExpress$_c$) × D$_s$	0.073***	0.101***	0.093***	0.096***	0.073***	0.071***	0.071***
	(0.017)	(0.026)	(0.015)	(0.014)	(0.017)	(0.017)	(0.017)
Post2010$_t$ × ln(AliExpress$_c$)	0.043***	0.033**	0.044***	0.035***		0.044***	0.044***
	(0.014)	(0.016)	(0.014)	(0.013)	(0.000)	(0.014)	(0.014)
Post2010$_t$ × D$_s$	0.408***	0.462***	0.557***	0.564***	0.406***	0.401***	
	(0.080)	(0.102)	(0.069)	(0.064)	(0.081)	(0.083)	(0.000)
Post2010$_t$ × ln(Distance$_c$)	−0.016	−0.016	−0.016	−0.016		−0.014	−0.015
	(0.018)	(0.022)	(0.018)	(0.018)	(0.000)	(0.020)	(0.019)
Post2010$_t$ × FTA$_{ct}$	0.037	0.017	0.037	0.037		0.037	0.036
	(0.030)	(0.038)	(0.030)	(0.030)	(0.000)	(0.031)	(0.031)
Post2010$_t$ × Language$_c$)	0.018	0.049	0.018	0.019		0.020	0.020
	(0.033)	(0.042)	(0.033)	(0.033)	(0.000)	(0.035)	(0.034)
Post2010$_t$ × Border$_c$)	0.034	0.083	0.034	0.035		0.040	0.039
	(0.054)	(0.063)	(0.054)	(0.054)	(0.000)	(0.057)	(0.055)
Post2010$_t$ × Legal$_c$)	0.015	−0.005	0.015	0.015		0.015	0.015
	(0.022)	(0.019)	(0.023)	(0.023)	(0.000)	(0.023)	(0.023)
D$_s$ × ln(AliExpress$_c$)	0.103**	−0.022	0.161***	0.148***	0.104**		0.105**
	(0.041)	(0.041)	(0.053)	(0.052)	(0.042)	(0.000)	(0.042)
ln(Internet$_{ct}$)	−0.035***	−0.029***	−0.035***	−0.035***		−0.038***	−0.038***
	(0.009)	(0.008)	(0.009)	(0.009)	(0.000)	(0.010)	(0.010)
ln(GDP$_{ct}$)	0.055	0.027	0.054	0.053		0.041	0.043
	(0.167)	(0.173)	(0.167)	(0.167)	(0.000)	(0.181)	(0.177)
ln(POP$_{ct}$)	−0.065	−0.073	−0.064	−0.064		−0.054	−0.055
	(0.201)	(0.211)	(0.202)	(0.202)	(0.000)	(0.220)	(0.212)
ln(GDPCap$_{ct}$)	−0.052	−0.074	−0.051	−0.050		−0.039	−0.043
	(0.171)	(0.176)	(0.171)	(0.171)	(0.000)	(0.185)	(0.180)
FTA$_{ct}$	−0.043**	−0.053**	−0.043**	−0.043**		−0.043**	−0.042**
	(0.018)	(0.023)	(0.018)	(0.018)	(0.000)	(0.019)	(0.019)
Post2010$_t$ × ln(Distance$_c$) × D$_s$		0.003					
		(0.031)					
Post2010$_t$ × FTA$_{ct}$ × D$_s$		0.057					
		(0.046)					
Post2010$_t$ × Language$_c$) × D$_s$		−0.092					
		(0.058)					
Post2010$_t$ × Border$_c$) × D$_s$		−0.137***					
		(0.048)					
Post2010$_t$ × Legal$_c$) × D$_s$		0.060*					
		(0.034)					
D$_s$ × ln(Distance$_c$)		−0.030					
		(0.056)					
D$_s$ × FTA$_{ct}$		0.030					
		(0.033)					
D$_s$ × Language$_c$		−0.026					
		(0.055)					
D$_s$ × Border$_c$		−0.090					
		(0.092)					
D$_s$ × Legal$_c$		0.039					
		(0.055)					
D$_s$ × ln(Internet$_{ct}$)		−0.024*					
		(0.014)					
D$_s$ × ln(GDP$_{ct}$)		0.034					
		(0.035)					
D$_s$ × ln(POP$_{ct}$)		0.074**					
		(0.031)					
D$_s$ × ln(GDPCap$_{ct}$)		0.108***					
		(0.040)					
FE	c,t,s	c,t,s	c,t,s	c,t,s	ct,s	cs,t	st,c
SE Cluster	c	c	c	c	c	c	c
N Countries	41	41	41	41	41	41	41
N Years	20	20	20	20	20	20	20
N Sectors	97	97	97	97	97	97	97
Observations	69,068	69,068	69,068	69,068	69,068	69,068	69,068
Adjusted R^2	0.490	0.503	0.498	0.499	0.490	0.837	0.541

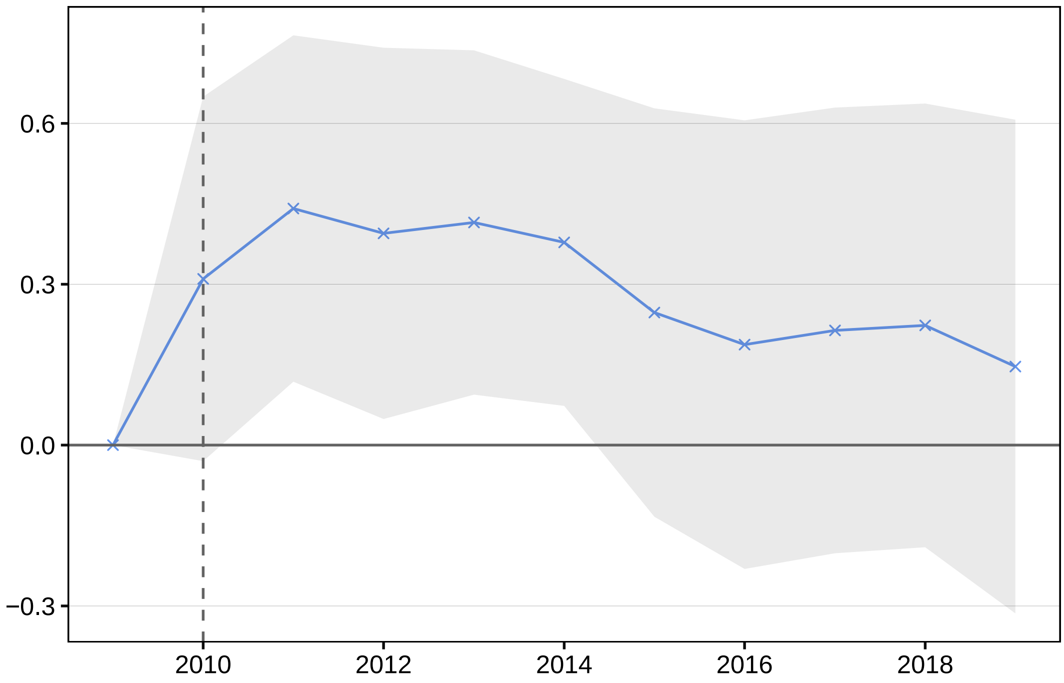
Notes: Table A3 reports robustness checks for the specification with more sectors. Standard errors culstered by countries are reported in parenthesis. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A4: AliExpress and Distance in a Gravity Setup

	Dependent variable:				
	$\ln(\text{Export}_{ijs})$				
	(1)	(2)	(3)	(4)	(5)
$\ln(\text{Distance}_{ij})$	-1.074^{***}	-1.094^{***}	-1.127^{***}	-0.898^{***}	-0.694^{***}
	(0.035)	(0.035)	(0.040)	(0.150)	(0.170)
$\ln(\text{Distance}_{ij}) \times \text{China}_i$		0.799^{***}	0.775^{***}	0.422	0.164
		(0.140)	(0.153)	(0.607)	(0.645)
$\ln(\text{Distance}_{ij}) \times \text{China}_i \times \text{D}_s$			0.416^{***}		0.612
			(0.127)		(0.744)
$\ln(\text{Distance}_{ij}) \times \text{China}_i \times \ln(\text{AliExpress}_j)$				-0.094	-0.144
				(0.140)	(0.158)
$\ln(\text{Distance}_{ij}) \times \text{China}_i \times \text{D}_s \times \ln(\text{AliExpress}_j)$					0.048
					(0.184)
Legal_{ij}	0.185^{***}	0.200^{***}	0.211^{***}	0.200^{***}	0.214^{***}
	(0.043)	(0.043)	(0.046)	(0.042)	(0.046)
Language_{ij}	0.210^{**}	0.203^{**}	0.253^{***}	0.193^{**}	0.231^{**}
	(0.086)	(0.085)	(0.095)	(0.087)	(0.097)
Border_{ij}	0.543^{***}	0.538^{***}	0.536^{***}	0.552^{***}	0.567^{***}
	(0.122)	(0.120)	(0.136)	(0.121)	(0.136)
FTA_{ij}	0.397^{***}	0.380^{***}	0.464^{***}	0.379^{***}	0.461^{***}
	(0.065)	(0.065)	(0.073)	(0.065)	(0.073)
Colony_{ij}	0.435^{***}	0.431^{***}	0.532^{***}	0.441^{***}	0.548^{***}
	(0.125)	(0.126)	(0.139)	(0.127)	(0.141)
FE	i,j	i,j	i,j,s	i,j	i,j,s
SE Cluster	ij	ij	ij	ij	ij
N Country Pairs	1795	1795	1795	1795	1795
N Countries	43	43	43	43	43
N Sectors	97	97	97	97	97
Observations	122,056	122,056	122,056	122,056	122,056
Adjusted R^2	0.327	0.327	0.529	0.327	0.529

Notes: Table A4 reports a gravity-type analysis of the AliExpress impact on Chinese exports. "ln(Export)" is the natural logarithm of trade flows (millions of US dollars) from country "i" to country "j". "ln(Distance)" is the natural logarithm of distance (Km) between country "i" and country "j". "China" is a dummy which equals one if country "i" is China. "Legal", "Language", "Border", "FTA" and "Colony" are dummy variables which equal one if one country in the country pair has, respectively, either a free trade agreement, a common official language, a common border, a common legal system or was a colony with the other country in the pair. The gravity sample includes the 42 countries for which a measure of the number of visitors on AliExpress.com is available plus China. All trade data (import at the 2-HS level) is from 2016. Standard errors culstered by country pairs are reported in parenthesis. *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Figure A3: Impact of AliExpress over time in the Apparel Sector (with respect to 2009)



Notes: Figure A3 reports the evolution of the AliExpress impact with respect to 2009 (year before the introduction of AliExpress) over time. The estimated coefficients here are from Equation 6. The shaded area is the 95% Confidence Interval of the estimated coefficient.

$$
\begin{aligned}
Export_{ct} = &\sum_{y=2010}^{2019} (\beta_{1y} 1\{y = t\} \times AliExpress_c) + \\
&\sum_{y=2010}^{2019} (\beta_{2y} 1\{y = t\} \times \boldsymbol{X}_c' \boldsymbol{\gamma}) + \\
&\lambda X_{ct} + \boldsymbol{\delta}_t + \boldsymbol{\delta}_c + \alpha + \epsilon_{ct}
\end{aligned} \tag{6}
$$