Duan, Jianhua; Xuefeng, Qian; Das, Kuntal K.; Meriluoto, Laura; Reed, W. Robert

Article

A replication of "The role of intermediaries in facilitating trade" (Journal of International Economics, 2011)

Economics: The Open-Access, Open-Assessment E-Journal

Provided in Cooperation with:
Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

Suggested Citation: Duan, Jianhua; Xuefeng, Qian; Das, Kuntal K.; Meriluoto, Laura; Reed, W. Robert (2020) : A replication of "The role of intermediaries in facilitating trade" (Journal of International Economics, 2011), Economics: The Open-Access, Open-Assessment E-Journal, ISSN 1864-6042, Kiel Institute for the World Economy (IfW), Kiel, Vol. 14, Iss. 2020-21, pp. 1-30, https://doi.org/10.5018/economics-ejournal.ja.2020-21

This Version is available at:
https://hdl.handle.net/10419/220003

Economics
The Open-Access, Open-Assessment E-Journal

Vol. 14, 2020-21 | June 22, 2020 | http://dx.doi.org/10.5018/economics-ejournal.ja.2020-21

A replication of "The role of intermediaries in facilitating trade" (Journal of International Economics, 2011)

Jianhua Duan, Xuefeng Qian, Kuntal K. Das, Laura Meriluoto, and W. Robert Reed

Abstract

This study replicates Ahn, Khandelwal, and Wei's (2011) model of intermediary trade. The study produces two main results. First, the authors are able to reproduce empirical evidence for AKW's three main predictions for Chinese exports. This is impressive because much of the data for their replication are independently sourced. However, when the authors subject their model to additional tests, they find that the evidence is not robust. Using more recently available data to test AKW´s first prediction, the authors estimate coefficients that are wrong-signed and significant. When they re-analyze the evidence supporting the second and third predictions, they find that the full sample results mask significant heterogeneity across Chinese regions. In many cases, key coefficients are insignificant. In a few cases, they are wrong-signed and significant. Finally, using multiple versions of a key variable measuring the number of required import documents by country, the authors discover that the results are not robust across versions.

The paper´s data set: https://doi.org/10.7910/DVN/BT8JPN

(Replication Study)

JEL F1
Keywords Intermediaries; exports; productivity; heterogeneous firms; China

Authors

Jianhua Duan, University of Canterbury, Christchurch, New Zealand
Xuefeng Qian, ✉ Zhongnan University of Economics and Law, Wuhan, China, xfqian@126.com
Kuntal K. Das, University of Canterbury, Christchurch, New Zealand
Laura Meriluoto, University of Canterbury, Christchurch, New Zealand
W. Robert Reed, Department of Economics and Finance, University of Canterbury, Christchurch, New Zealand

Citation Jianhua Duan, Xuefeng Qian, Kuntal K. Das, Laura Meriluoto, and W. Robert Reed (2020). A replication of "The role of intermediaries in facilitating trade" (Journal of International Economics, 2011). *Economics: The Open-Access, Open-Assessment E-Journal*, 14 (2020-21): 1–30. http://dx.doi.org/10.5018/economics- ejournal.ja.2020-21

Received October 10, 2019 **Published as Economics Discussion Paper** November 19, 2011
Revised April 16, 2020 **Accepted** April 23 2020 **Published** June 22, 2020

1 Introduction

In her widely acclaimed book, *Statistical Inference as Severe Testing*, Deborah Mayo (2018) argues that a root cause of science's reproducibility problem is that scientific claims are not strongly tested. She advocates for "severe testing", where "A claim is severely tested to the extent it has been subjected to and passes a test that probably would have found flaws, were they present" (Mayo, 2018, p. xii). Replication can play that role both by assessing the evidence in the original study, and subjecting its claims to strong tests.

This paper replicates Ahn, Khandelwal, and Wei (AKW), "The role of intermediaries in facilitating trade" (*Journal of International Economics*, 2011). AKW build on the seminal work of Melitz (2003) linking export behaviour to heterogeneous firm productivity. Whereas the exporting decision in Melitz (2003) consists of to-export or not-to-export, AKW build in a third option, exporting through intermediaries.

AKW is important because it provides both a theoretical foundation for the existence of trade intermediaries, and supporting empirical evidence. While other theoretical models of intermediation exist (cf. Akerman, 2018; Poncet and Xu, 2018; Blum, Claro, and Horstmann, 2018; and Chan, 2018), AKW was the first to be published in a major journal and is the most cited. At the time of this writing, it has been cited 153 times in Scopus, and 596 times in Google Scholar.

Two main results follow from our replication. First, we are able to reproduce the empirical evidence that AKW provide in support of their model. This is impressive because we had to independently source much of the data used by AKW. However, when the data are re-analysed and updated, we find that empirical support for their model is substantially diminished.

We note that we see our role as independent auditors, neither predisposed to produce evidence in favor for or against AKW's model. One of the advantages of submitting our replication research to *Economics: The Open Access, Open Assessment E-Journal* is that the journal states that it "will publish both confirmations and disconfirmations of original studies. The only consideration will be quality of the replicating study."[1] Thus we do not have an incentive to cherry-pick evidence in support of a given result.

We also note that we have consulted with the original authors: JaeBin Ahn, Amit Khandelwal and Shang-Jin Wei. AKW were very gracious in sharing their data and code with us. In addition, they patiently answered our queries and gave us comments on our research. While we have some disagreements with them regarding the interpretation of our results, we try our best to give voice to their concerns in our discussion below. They will have the opportunity to raise additional concerns in a reply.

Our study proceeds as follows. Section 2 provides the empirical and theoretical context for AKW's research on intermediation. Section 3 briefly summarizes AKW's model of trade intermediation so that the reader can understand how their empirical predictions derive from the model. Section 4 describes our data.[2] Section 5 reports our attempt to reproduce the key

[1] This statement is taken from the journal's website: http://www.economics-ejournal.org/special-areas/replications-1.

[2] All the data and code to reproduce the results in our paper is publicly available at Harvard's Dataverse: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/BT8JPN.

evidence provided by AKW in support of their model. Sections 6 through 9 present and discuss additional test results from re-analysing and extending AKW's data. Section 10 concludes.

2 Empirical and theoretical context

Intermediation is widely recognized as playing a prominent role in world trade markets. In the early 1980's, 300 Japanese traders (non-manufacturing firms) accounted for 80% of Japan's trade (Rossman, 1984). Spulber (1996) documents that in 1995, intermediaries accounted for about a quarter of the GDP for the U.S., and close to two million firms operated in the U.S. intermediation industry. In 2002, American intermediaries accounted for 44% and 56% of exporting and importing firms, and 11% and 24% of export and import value, respectively (Bernard, Jensen, Redding, and Schott, 2010). The economies of Hong Kong, Singapore, and Dubai have developed specialised expertise in entrepôt trade, which has greatly contributed to their economic development (Feenstra and Hanson, 2004). For these reasons and more, intermediation has received much empirical attention (Antràs and Costinot, 2011; Blum, Claro, and Horstmann, 2018; Qu, Raff, and Schmitt, 2014; Tang and Zhang, 2012).

The literature on intermediaries in international trade makes two important observations. First, a significant fraction of international trade is channelled through intermediaries. Second, there are systematic variations in the mode of exports not only across firms within an industry, but also across industries. Fundamental to the role of intermediaries is connecting buyers and sellers worldwide.

A variety of reasons has been suggested for why intermediaries exist. One reason is that they mitigate uncertainty regarding demand and supply, or when buyer and seller characteristics are unobservable (Spulber, 1996). Another reason is that trade intermediaries perform quality assurance. Uncertainty about product quality creates the familiar problems of adverse selection and moral hazard. Trade intermediaries can alleviate this problem by screening the quality of the products and revealing it to the customers (Dasgupta and Mondria, 2018). Intermediaries can also facilitate export activity by providing a credit-constrained firm a channel through which capital market frictions are reduced, thus enhancing the gains from trade (Chan, 2018).

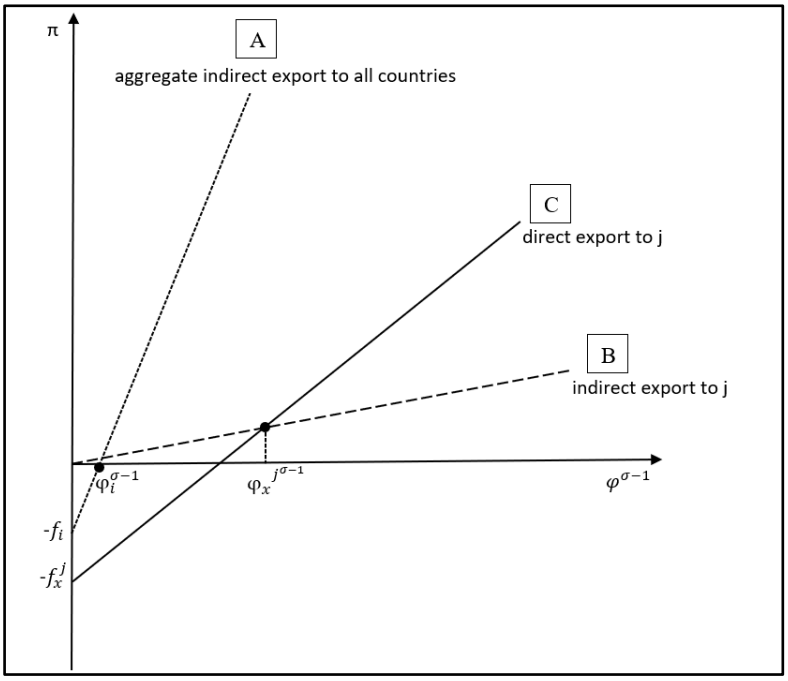
Intermediaries are not always seen as beneficial. Antràs and Costinot (2011) present a model with search frictions where intermediaries provide market access to farmers. They show that, depending on the kind of market integration being considered, intermediation can lower the welfare of farmers in developing countries. Similarly, Sheveleva and Krishna (2017) show how a hold-up problem between farmers and intermediaries, arising from contractual incompleteness, leads to the nonexistence of markets for certain agricultural goods.

While much of the literature has focussed on intermediaries as facilitating international matches, AKW argue that intermediaries exist primarily to overcome trade costs. Part of the appeal of AKW's model is that it makes three straightforward predictions that are strongly supported in their empirical analysis.

3 AKW's model of trade intermediation and three predictions

The model. Figure 1 represents the main ideas underlying AKW's theory of intermediated trade. The vertical axis measures profits from exporting (π). The horizontal axis measures firm productivity ($\varphi^{\sigma-1}$).[3] Note that $\pi = 0$ if the firm does not export, and coincides with the horizontal axis. The steep dotted line (A) identifies the profit-productivity nexus for the firm if it sells its output to an intermediary that then exports it overseas. The initial cost of contracting with the intermediary is f_i. The intermediary then repackages the firm's output and sells it in *N* overseas markets. The associated increase in profits resulting from an increase in productivity is given by the slope of A. The point at which A crosses the horizontal axis identifies the minimal productivity ($\varphi_i^{\sigma-1}$) at which the firm will switch from selling only in the domestic market to exporting via an intermediary.

Figure 1: AKW's model of intermediary trade



Notes: The vertical axis measures profits from exporting (π). The horizontal axis measures firm productivity ($\varphi^{\sigma-1}$). The horizontal axis represents the profits from exporting if the firm does not export (=0). Line A represents the profits to the firm if it sells its output to an intermediary that then repackages the output and sells it in *N* overseas markets. The point at which A crosses the horizontal axis identifies the minimal productivity ($\varphi_i^{\sigma-1}$) at which the firm will switch from selling only in the domestic market to also exporting via an intermediary. Line B represents the profit-productivity nexus for indirectly exporting to country *j* through an intermediary. Line C indicates the associated profits for directly exporting to *j*. The productivity where the two lines intersect ($\varphi_x^{j\sigma-1}$) identifies the threshold where a firm transitions from indirectly selling its product to country *j* to directly selling to *j*.

[3] Mathematical notation is taken from Ahn, Khandelwal, and Wei (2011). The interested reader is referred there for further detail.

A firm can choose to directly export to a specific country *j* even while selling to other countries via an intermediary. This decision is represented by lines B and C. Line B represents the profit-productivity nexus for indirectly exporting to country *j* through an intermediary. Since the firm has already contracted with the intermediary, it does not need to pay any additional fixed cost for (indirect) exports to *j*. B is flatter than A because A captures profits from indirect exports to all countries, while B only captures profits from country *j*.

If, instead, the firm chooses to directly export to *j*, it must bear all the fixed costs of entering that market and finding buyers. This is represented by f_x^j. The reward to doing so is that it enjoys higher marginal profits. The intermediary must repackage and rebrand the product, which raises marginal costs, and, hence, price, in the foreign market. This higher price decreases the quantity demanded, which results in lower profits for the firm. The firm that directly exports its product avoids these costs. This enables it to sell its product at a lower price overseas, which increases profits. As a result, the slope of line C is greater than the slope of line B. The productivity where the two lines intersect ($\varphi_x^{j\,\sigma-1}$) identifies the threshold where a firm transitions from indirectly selling its product to country *j* to directly selling to *j*.

Three predictions. AKW use the preceding model to derive the following three predictions:

> Prediction #1: "…we would expect a hockey stick relationship between productivity and direct exports ‑ only high productivity firms directly export while low and intermediate productivity firms do not ‑ and an inverted U-shape relationship with indirect exports" (page 78).

As seen in Figure 1, the model predicts that increases in productivity will have no effect on direct exports for low productivity firms. Increases in productivity for high productivity firms will generate linear increases in profits, giving a "hockey stick" relationship. In contrast, increases in productivity for low productivity firms will first generate an increase in indirect exports as firms transition from no exports to exporting via an intermediary. However, further increases in productivity are predicted to decrease indirect exports as firms transition from exporting via an intermediary to directly exporting, giving an inverted U-shape relationship.

> Prediction #2: "Exports by intermediaries should be more expensive than direct exporters" (page 79).

This prediction follows directly from the following assumption of the model: "Intermediaries purchase varieties from manufacturers at the same price as domestic consumers (there is no price discrimination) and incur an additional marginal cost of selling these varieties abroad. This additional marginal cost captures re-labeling, packaging and other per-unit costs associated with taking the title of varieties from the manufacturers. The price of indirectly exported varieties is therefore higher than the price of directly exported varieties by this factor" (page 75).

> Prediction #3: "… the share of exports through intermediaries is larger in countries with (i) smaller market size, (ii) higher variable trade costs, or (iii) higher fixed costs of exporting" (page 76).

This prediction also follows from Figure 1. Ceteris paribus, smaller destination markets or higher variable trade costs cause the slopes of lines C and B to swing down. However, for any positive productivity level, line C will swing down by more than line B. The result is that the threshold from switching from indirect to direct exporting ($\varphi_x^{j\sigma-1}$) moves to the right, thus increasing the share of exports through intermediaries. Higher fixed costs associated with exporting shift line C down, without affecting line B. This again causes $\varphi_x^{j\sigma-1}$ to move right, increasing the share of indirect exports.

4 Data

Most of the data used to test AKW's three predictions are drawn from two sources: China Customs Data and Enterprise Survey Data (Chinese firms only). The customs data record international trade information for individual Chinese firms over the 2000–2005 period. They contain detailed information for each firm-product-partner transaction, including information on product type, country destination, and the price and quantity of the transactions. The extensive coverage of the China Customs Data allows a relatively complete record of trade transactions at the country level. The China Customs Data also records information about the firm, including its name. AKW exploit a Chinese naming convention to identify intermediary firms. As they describe it,

> "We identify the set of intermediary firms based on Chinese characters that have the English-equivalent meaning of "importer", "exporter", and/or "trading" in the firm's name. A useful feature about firm names in China is that they are often very descriptive (a convention that might be traced to a time when the country was under central planning and the planners favored descriptive company names). Many firms founded during the post-1980 reform era continue to adopt this naming convention" (pages 76f.).

While AKW's approach allows them to distinguish intermediary firms from other firms in the China Customs Data ("direct firms"), it does not allow them to identify how much of direct firms' output is sold to intermediaries. A key aspect of AKW's model is the decision by firms to directly export their product overseas versus indirectly exporting through intermediaries. A core prediction of the model relates how firm productivity affects the share of total sales going to direct or indirect exports. To address this issue, AKW turn to the World Bank's Enterprise Surveys.

The Enterprise Survey Data use "standard survey instruments to collect firm-level data on the business environment from business owners and top managers".[4] The survey covers a wide range of subject areas, and includes data from many countries. A major advantage of the Enterprise Survey Data is that it reports the proportions of an establishment's sales exported (i) directly and (ii) through a distributor. The survey also records various measures of productivity. Survey data are available for the years 2002/2003 and 2012, the last of which was unavailable to AKW at the time of their writing.

[4] Taken from https://www.enterprisesurveys.org/portal/login.aspx (accessed on 9 January 2019).

To test Prediction #3, AKW employ a number of different variables to capture market characteristics. Distance from trading partner, Most Favored Nation (MFN) tariff, size of ethnic Chinese population in the destination country, and market size (measured by GDP) are hypothesized to proxy trade costs. Increases in the first two are assumed to raise variable costs, and increases in the latter two are assumed to lower costs. Importing procedures, measured as the number of documents required for importing, are assumed to be positively associated with fixed costs.

AKW graciously shared their programming code to assist us in our replication project. They were unable to share their China Customs Data with us because it is proprietary. When we encountered difficulties or had questions, AKW assisted us by providing their list of intermediary firms, their data on importing procedures, and the MFN tariff data. For the remainder of the data, we collected them ourselves. The Appendix reports the variables used in our reproduction of AKW, along with the respective data sources.

The fact that we had to independently collect much of the data ourselves inevitably meant that discrepancies would arise, not the least because some data are updated over time. Nevertheless, as we show below, our efforts to reconstruct their data resulted in us being able to closely reproduce AKW's reported estimates.

5 Reproduction of AKW´s results

The main results from testing AKW's three predictions are reported in Tables 4–6 of their paper. Prediction #1 states that the relationship between productivity and direct exports should have a "hockey stick" shape, while that between productivity and indirect exports should show an inverted U-shape. Table 1 reports the reproduction of AKW's test of this prediction.

Panel A shows the results of separately regressing the share of total exports due to direct exports on three different measures of productivity: sales, employment and sales per worker ("labor productivity"). Columns (1.a), (2.a), and (3.a) copy the results from their paper. All of the linear terms are positive. None of the quadratic terms are significant. The linear productivity variable is significant at the 5 percent level in (3.a) and at the 10 percent level in (2.a). AKW interpret the linear relationship between productivity and direct export share as consistent with the "hockey stick" prediction from Prediction #1.

Using the Enterprise Survey Data that we downloaded from the World Bank online data site, we re-estimated AKW's specifications. Our reproduction results are reported in Columns (1.b), (2.b), and (3.b). In each case, we are able to exactly reproduce their results. The only discrepancy between our results and AKW's is seen in the R-squared and Observation values for the second and third specifications, which are reversed. We suspect this is due to a typographical error in the published version of their manuscript.

To facilitate interpretation of the many results to follow, we highlight estimates that do not provide support for AKW's theory. In this case, two of the three linear coefficients, while

Table 1. Reproduction of AKW's results: Prediction #1

A. Direct export share

Variables	AKW (1.a)	Reproduction (1.b)	AKW (2.a)	Reproduction (2.b)	AKW (3.a)	Reproduction (3.b)
{Log Sales}	0.015	0.015	----	----	----	----
	(0.013)	(0.013)				
{Log Sales}^{2}	0.0010	0.0010	----	----	----	----
	(0.0007)	(0.0007)				
{Log Employment}	----	----	0.041*	0.041*	----	----
			(0.024)	(0.024)		
{Log Employment}^{2}	----	----	0.001	0.001	----	----
			(0.002)	(0.002)		
{Log Labor Productivity}	----	----	----	----	0.024**	0.024**
					(0.010)	(0.010)
{Log Labor Productivity}^{2}	----	----	----	----	0.001	0.001
					(0.001)	(0.001)
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.12	0.12	0.08	0.11	0.11	0.08
Observations	2469	2469	2340	2364	2364	2340

B. Indirect export share

Variables	AKW (1.a)	Reproduction (1.b)	AKW (2.a)	Reproduction (2.b)	AKW (3.a)	Reproduction (3.b)
{Log Sales}	0.034***	0.034***	----	----	----	----
	(0.009)	(0.009)				
{Log Sales}^{2}	-0.002***	-0.002***	----	----	----	----
	(0.000)	(0.000)				
{Log Employment}	----	----	0.039**	0.039**	----	----
			(0.016)	(0.016)		
{Log Employment}^{2}	----	----	-0.003**	-0.003**	----	----
			(0.001)	(0.001)		
{Log Labor Productivity}	----	----	----	----	0.016**	0.016**
					(0.007)	(0.007)
{Log Labor Productivity}^{2}	----	----	----	----	-0.001*	-0.001*
					(0.001)	(0.001)
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.05	0.05	0.05	0.05	0.05	0.05
Observations	2570	2570	2437	2461	2461	2437

NOTES**:** Columns (1.a), (2.a), and (3.a) report the results from AKW's Table 4. The results from independently reproducing their results are reported in Columns (1.b), (2.b), and (3.b), respectively. Data are taken from the 2002 and 2003 releases of the World Bank's Enterprise Survey Data. The dependent variable is direct exports (Panel A) / indirect exports (Panel B) as a share of total sales. Regressions are estimated using OLS. In addition to the respective productivity variables, all specifications include industry fixed effects. Standard errors are reported in parentheses below coefficient estimates. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively. The gray-shaded cells indicate that the respective estimates have the predicted sign, but are not statistically significant at the 5-percent level.

having the predicted sign, are insignificant at the 5 percent level. We shade the respective cells in gray to indicate that they, being insignificant, do not provide support for AKW's predictions even though they have the right sign. Subsequent analysis will also make use of two other color codes. A rose-shaded cell indicates that the respective coefficient is wrong-signed and insignificant at the 5-percent level. A red-shaded cell indicates that it is wrong-signed and significant. Coefficients that have the predicted sign and are significant are unshaded.

Panel B repeats the analysis, except that the dependent variable is now the share of total exports due to indirect exporting. An inverted U-shape implies that the linear productivity variable should have a positive coefficient, with the coefficient for the quadratic term being negative. AKW's results confirm this prediction for both the linear and quadratic terms, with all being statistically significant at the 5 percent level except the quadratic labor-productivity coefficient. Once again, our reproduction of AKW's analysis produces identical results to theirs. Taken together, the results from Table 1 generally confirm AKW's first prediction. Of the nine coefficients, all have the predicted sign and six are statistically significant at the 5 percent level.

AKW's second prediction is that unit value, the average export price of a good, is greater for intermediary firms than for other exporting firms. To test this prediction, AKW use the China Customs Data from 2005 and regress log unit value on a dummy variable indicating that the given firm is an intermediary. The extensive coverage of the China Customs Data, and the fact that the data cover individual transactions, ensures that there are a large number of observations. Table 2 reports the results.

AKW test three specifications. The first controls for product category and ownership type (1.a). The second adds controls for firm size (2.a), and the third adds controls for destination country (3.a). For all three specifications, their estimated coefficients are positive and significant at the 1 percent level, consistent with their prediction. Our reproduction results are reported in Columns (1.b) to (3.b), respectively. For reasons that are unclear, our China Customs Data produce a substantially larger number of observations than AKW's. Nevertheless, in every case our estimated coefficient for the intermediary firm dummy is positive and

Table 2: Reproduction of AKW's results: Prediction #2

Variables	**AKW (1.a)**	**Reproduction (1.b)**	**AKW (2.a)**	**Reproduction (2.b)**	**AKW (3.a)**	**Reproduction (3.b)**
{Intermediary}$_f$	0.067***	0.064***	0.051***	0.045***	0.023***	0.017***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Quartic firm size controls	No	No	Yes	Yes	Yes	Yes
Fixed effects	po	po	po	po	cpo	cpo
R-squared	0.79	0.79	0.79	0.79	0.85	0.85
Observations	4,594,598	5,193,328	4,594,598	5,193,328	4,594,598	5,193,328

NOTES**:** Columns (1.a), (2.a), and (3.a) report the results from Columns (1) – (3) of AKW's Table 5. The results from independently reproducing their results are reported in Columns (1.b), (2.b), and (3.b), respectively. The dependent variable is the log unit value from individual transactions. Regressions are estimated using OLS with cluster robust standard errors grouped by product (HS8). "p", "o", and "c" refer to paired/triplet fixed effects based on product, ownership, and country. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively.

Table 3: Reproduction of AKW's results: Prediction #3

Variables	AKW (1.a)	Reproduction (1.b)	AKW (2.a)	Reproduction (2.b)	AKW (3.a)	Reproduction (3.b)	AKW (4.a)	Reproduction (4.b)
{Log Distance}$_c$	0.032***	0.027***	0.026***	0.029***	0.025***	0.031***	0.025***	0.030***
	(0.008)	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.009)
{Log GDP}$_c$	-0.022***	-0.021***	-0.021***	-0.019***	-0.019***	-0.019***	-0.019***	-0.017***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
{Log Chinese Population}$_c$	----	----	-0.002*	-0.003**	-0.004*	-0.003***	-0.004*	-0.004***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
{# of Importing Procs}$_c$	----	----	----	----	0.003**	0.004***	0.003**	0.004***
					(0.001)	(0.001)	(0.001)	(0.001)
{MFN Tariff}$_{hc}$	----	----	----	----	----	----	0.059**	0.034*
							(0.022)	(0.020)
HS6 FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.15	0.15	0.16	0.16	0.17	0.17	0.18	0.17
Observations	267,201	264,662	221,373	227,304	207,594	212,901	185,975	188,120

NOTES**:** Columns (1.a), (2.a), (3.a), and (4.a) report the results from Columns (1) – (4) of AKW's Table 6. The results from independently reproducing their results are reported in Columns (1.b), (2.b), (3.b), and (4.b), respectively. The dependent variable in each regression is the share of intermediary exports of total country HS6 exports. Regressions are estimated using OLS with cluster robust standard errors grouped by country. All regressions also include fixed effects for product type (HS6). *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively. The gray-shaded cells indicate that the respective estimates have the predicted sign, but are not statistically significant at the 5-percent level

significant at the 1 percent level, and relatively close to AKW's estimated values. These results strongly confirm AKW's second prediction.

AKW refer to their last prediction (Prediction #3) as "the central prediction of the model: intermediary shares will be systematically correlated with destination market characteristics" (page 79). AKW focus on five destination market characteristics: distance, market size (as measured by GDP), size of the ethnic Chinese population, number of "required documents for imports", and MFN tariff on Chinese HS6 products. As discussed above, the expected signs of these coefficients are positive, negative, negative, positive, and positive, respectively.

Table 3 reports the results. AKW estimate four specifications. The first specification includes distance and GDP. The next specification adds size of Chinese population in the destination country. The third adds number of required documents, and the last adds tariffs. Their results are shown in Columns (1.a) to (4.a). The dependent variable is the share of total country HS6 exports due to intermediary exports. Aggregating to the country-HS6 product level explains the reduction in observations from Table 2 to Table 3.

All of the coefficients have the expected signs. Of the 14 estimated coefficients, only the coefficients for Chinese population are not significant at the 5-percent level (see gray-shaded cells). Our reproduction estimates, which are based on our independently sourced China Customs Data, though we use AKW's required documents and tariff data, provide, if anything, even stronger support for AKW's third prediction. All of the 14 estimated coefficients have the expected signs, and only one – the coefficient for the tariff variable – is not significant at the 5 percent level.

In conclusion, using data that are largely independently sourced, we confirm the main results that AKW report in support of the three predictions derived from their model. It speaks highly of AKW's careful handling of their data that we can do this. In the next section, we re-analyze and extend their data to subject it to further testing.

6 Robustness check: Prediction #1

Description. AKW's primary test of Prediction #1 relies on Enterprise Survey Data from years 2002 and 2003. However, after AKW published their paper in 2011, the World Bank released another round of Enterprise Survey Data in 2012. Our first set of additional tests consists of repeating AKW's analysis with these new data.

Results. The results for this robustness check are presented in Table 4. Panel A reports the results when the dependent variable is the share of total exports that are direct exports. To facilitate interpretation, the second column indicates the sign prediction from AKW's theory. AKW predict a "hockey shape" relationship between productivity and direct exports. This implies a positive linear relationship. They do not make a prediction about the quadratic term, so we denote that as indeterminate ("Indet"). Of the three linear productivity coefficients, two are positive and significant at the 5 percent level (Log Sales and Log Employment) and one is positive but insignificant (Log Labor Productivity). We gray-shade the latter cell to indicate that it does not provide supporting evidence for AKW's first prediction, though the respective coefficient does have the expected sign.

Table 4: Robustness check of AKW's results: Prediction #1

A. Direct export share

Variables	Expected Sign	2012 (1)	2012 (2)	2012 (3)
{Log Sales}	Pos	0.092** (0.038)	----	----
{Log Sales}^{2}	Indet	-0.002* (0.001)	----	----
{Log Employment}	Pos	----	0.050** (0.020)	----
{Log Employment}^{2}	Indet	----	-0.002 (0.002)	----
{Log Labor Productivity}	Pos	----	----	0.050 (0.068)
{Log Labor Productivity}^{2}	Indet	----	----	-0.002 (0.003)
Industry FEs	----	Yes	Yes	Yes
R-squared	----	0.06	0.07	0.03
Observations	----	1692	1691	1691

B. Indirect export share

Variables	Expected Sign	2012 (1)	2012 (2)	2012 (3)
{Log Sales}	Pos	-0.010 (0.028)	----	----
{Log Sales}^{2}	Neg	0.000 (0.001)	----	----
{Log Employment}	Pos	----	0.041*** (0.014)	----
{Log Employment}^{2}	Neg	----	-0.004*** (0.001)	----
{Log Labor Productivity}	Pos	----	----	-0.112** (0.048)
{Log Labor Productivity}^{2}	Neg	----	----	0.004** (0.002)
Industry FEs	----	Yes	Yes	Yes
R-squared	----	0.04	0.05	0.05
Observations	----	1692	1691	1691

NOTES: Columns (1), (2), and (3) re-estimate the regressions from Table 1 using data from the 2012 release of the World Bank's Enterprise Survey Data. The dependent variable is direct exports (Panel A) / indirect exports (Panel B) as a share of total sales. Regressions are estimated using OLS. In addition to the respective productivity variables, all specifications include industry fixed effects. Standard errors are reported in parentheses below coefficient estimates. "Pos", "Neg", and "Indet" refer to positive, negative, and indeterminate. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively.

The red-shaded cells indicate that the respective estimates have the wrong sign and are statistically significant at the 5-percent level. The rose-shaded cells indicate that the respective estimates have the wrong sign, but are not statistically significant at the 5-percent level. The gray-shaded cells indicate that the respective estimates have the predicted sign, but are not statistically significant at the 5-percent level.

Panel B repeats the analysis using indirect exports as the dependent variable. The prediction of an inverted U-shape relationship between productivity and indirect exports generates six sign predictions: a positive sign for the linear term and a negative sign for the quadratic term, for each of the three measures of productivity. Two of the six have the predicted sign (the linear and quadratic terms for Log Employment). Two of the six have the wrong signs but are insignificant and thus are shaded rose (Log Sales), and two of the six have wrong signs and are significant at the 5 percent level (Log Labor Productivity).

Summary. There are a total of nine coefficients that allow us to test AKW's Prediction #1 using the 2012 Enterprise Survey Data. Four of the nine are statistically significant and provide confirmation for their theory. Two of the nine are statistically significant and contradict their theory. The remaining three are statistically insignificant.

One of the advantages of the journal *Economics: The Open Access, Open Assessment E-Journal* is that readers have access to the reviewers' reports. Reviewer #1 notes that sales and employment are primarily concerned with size, so that labor productivity provides a better test of AKW's Prediction #1. Accordingly, the labor productivity estimates should receive greater weight. If one takes that position, the results are more damaging to AKW's prediction, because the associated coefficient in the direct export share regression is insignificant (though correctly signed), while both of the corresponding coefficients in the indirect export share regression are wrong-signed and significant.

Qualifications. The Enterprise Survey targets small, medium, and large companies in the non-agricultural, formal, private economy. It is noteworthy that both the 2002/2003 and 2012 surveys included relatively few firms compared to the total number of Chinese companies. The samples in Tables 1 and 4 consisted of 2340/2570 and 1691/1692 companies, respectively. The relatively small sizes of the surveys raise questions about their representativeness. The contrary results from the 2012 survey could be due to spurious correlation with nonrepresentative factors. However, if this is the case, it suggests that the positive results from the 2002/2003 survey may also be unreliable, providing coincidental, albeit spurious evidence for AKW's theory.

7 Robustness check: Prediction #2

Description. Table 5 provides both a re-analysis and an extension of AKW's test of Prediction #2, reported in Table 2 above. The theory predicts that intermediaries will sell their product at a higher price than firms that directly sell their product overseas. AKW use the 2005 China Customs Data to test that prediction. They use three specifications that cumulatively add more control variables. First, they include fixed effects to control for product and ownership type (Column 1.a in Table 2). Then they add in controls for firm size (Column 2.a in Table 2). Finally, they add in controls for destination country (Column 3.a in Table 2). We work with the final specification that includes the most controls.

Results. Our first approach consists of breaking up the full sample of China Customs Data observations into three, mutually exclusive, geographical regions: East, Central, and West. The original sample in our reproduction of AKW's results (cf. Column 3.b in Table 2) had over 5

Table 5: Robustness check of AKW's results: Prediction #2

A. subsamples by region and year

Sample	{Intermediary}$_f$ (Expect = Pos)	Observations
2005/East	0.022*** (0.004)	4,815,809
2005/Central	-0.058*** (0.013)	224,526
2005/West	-0.057*** (0.013)	166,825
2000	-0.003 (0.004)	1,731,580
2001	0.026*** (0.004)	2,019,600
2002	0.022*** (0.004)	2,688,711
2003	-0.023*** (0.003)	3,358,635
2004	-0.003 (0.003)	4,166,559

B. Further Analysis

Sample	{Intermediary}$_f$ (Expect = Pos)	Observations
2005/East1	-0.041*** (0.006)	638,145
2005/East2	-0.021*** (0.005)	1,890,221
2005/East3	0.039*** (0.003)	1,545,156
Panel (With Year FEs)	-0.126*** (0.008)	19,158,413

NOTES**:** The first three rows of Panel A ("2005/East", "2005/Central", "2005/West") re-estimate specification (3.a/3.b) from Table 2 using the same data, except that the full sample is divided into three subsamples. "East" includes Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Hainan. "Central" includes Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hunan, Hubei. "West" includes Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang, Inner Mongolia, Guangxi. The bottom five rows of Panel A ("2000"-"2004") re-estimate specification (3.a/3.b) from Table 2 using earlier years of the China Customs Data. The first three rows of Panel B ("2005/East1", "2005/East2", "2005/East3") further subdivide the top row of Panel A into three subsamples: "East1" is the Bohai Bay Economic Rim, including Beijing, Tianjin, Hebei, Liaoning, Shandong; "East2" is the Yangtze River Delta Economic Zone, including Shanghai, Jiangsu, Zhejiang; and "East3" is the Pearl River Delta Economic Zone, including Guangdong.

The dependent variable is the log unit value from individual transactions. Regressions are estimated using OLS with cluster robust standard errors grouped by product (HS8). In addition to the intermediary dummy variable, all regressions include triplet fixed effects for product, ownership, and country. "Pos" refers to positive. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively. The red-shaded cells indicate that the respective estimates have the wrong sign and are statistically significant at the 5-percent level. The rose-shaded cells indicate that the respective estimates have the wrong sign, but are not statistically significant at the 5-percent level.

million observations. The three subsamples have 4,815,809; 224,526; and 166,825 observations, respectively. Given the large number of observations in each of the subsamples, loss of statistical power from dividing the full sample is not a concern.5 As there is nothing in AKW's theory that suggests their prediction about higher unit prices depends on geographical location within China, breaking up the full sample into three subsamples has the advantage of providing three tests of their prediction rather than one.

The first three rows of Table 5/Panel A report the results. The results from firms located in China's eastern provinces provide support for AKW's second prediction, but the results from the central and western provinces do not. Both of the latter intermediary dummy coefficients are negative and significant at the 1 percent level. To indicate that these results provide evidence against AKW's theory, we color-code the associated cells in the table with red.

Our second set of tests consists of extending the analysis of AKW's model to earlier years of the China Customs Data. As noted by AKW, while they had access to China Customs Data for the years 2000 to 2005, they chose to focus on 2005 due to governmental restrictions on trading:

> "Another issue that could potentially complicate our analysis is that the Chinese government issued trading licenses for certain products prior to China's entry into the WTO. The WTO mandated that China liberalize the scope and availability of licenses so that within three years after accession, all enterprises would have the right to trade products without licenses. China's WTO accession document indicates that in the first year of accession, only wholly Chinese-invested firms with registered capital exceeding RMB 5 million could obtain direct trading rights. In the second year after accession, the minimum capital requirement for direct trading was RMB 3 million, and this fell to RMB 1 million by 2004. However, data from the World Bank's Enterprise Survey for China that covers 2002 and 2003 indicate that firms below this cutoff reported direct exports. This could be because export licenses were only required for a limited set of products and/or because these cutoffs were not stringently applied, at least for exports. By 2005, any firm that wished to directly trade with foreign partners was free to do so. So while we are confident that the licenses will not affect the interpretation of our results, the main analysis uses data for 2005 when the licenses had been removed" (AKW, page 77).

In other words, prior to 2005, firms were required to have government-issued licenses in order to export. There is evidence that this requirement was not strictly enforced. Further, it only applied to a subset of products.[6] Thus, AKW are comfortable using the 2002/2003 Enterprise Survey Data. Nevertheless, export data after 2004 is preferable because by 2005 firms were free to export without license restrictions.

While we recognize that the pre-2005 data is inferior to the 2005 data, we believe there is value in investigating these earlier data. We base this on the fact that AKW also use data from

---

5 Of the 4,815,809 observations in the "East" subsample, 43.6% involve transactions by intermediaries. Of the 224,526 observations in the "Central" subsample, 54.0% involve transactions by intermediaries. Of the 166,825 observations in the "West" subsample, 53.3% involve transactions by intermediaries.

6 AKW note in Footnote 17, page 78: "While there were some restrictions of trading during this period, they were limited to only a subset of products."

these earlier years as evidence in support of their theory (see Table 2 of AKW). Expanding the analysis to test Prediction #2 using pre-2005 data allows additional tests of AKW's model.

The last five rows of Table 5, Panel A report the results. There are five coefficients corresponding to the intermediary dummy variable for the year 2000 to 2004 regressions. Two provide support for AKW's prediction, being positive and significant (2001 and 2002). Of the remaining three, two are wrong-signed and insignificant (2000 and 2004), and one is wrong-signed and significant at the 1 percent level (2003).

We again refer to the reports of the original authors and reviewers that are available online at the *Economics E-Journal* website. The general consensus is that the geographical subsamples of Panel A were not persuasive because the Central and West estimates were based on much smaller sample sizes. Further, one could argue that the Central and West regions are subject to idiosyncratic influences as a result of their inland location. Accordingly, we followed the recommendation of one of the reviewers and further divided the Eastern provinces into three subsamples (East1, East2, East3).[7] The first three rows of Panel B report the results.

While the East3 estimate is consistent with AKW's Prediction #2, the estimates for East1 and East2 are wrong-signed and statistically significant. Following up on another suggestion by a reviewer, we also combined the 2000–2005 data and estimated a two-way, fixed effects model (firm + year). This also produced a wrong-signed and statistically significant estimate.

Summary. Table 5 provides twelve "additional" tests of AKW's second prediction. "Additional" is in quotation because the respective 2005 subsamples are not independent of the full 2005 sample reported in Table 2. Nevertheless, only four of the twelve estimates confirm AKW's model. Six of the twelve estimates contradict AKW's model and are significant at the 1 percent level. The remaining two are inconclusive because they are statistically insignificant.

Qualifications. One concern with our robustness checks relates to our decision to divide the sample into geographical subsample. To address concerns of cherry-picking results, we should have completed a pre-analysis plan. We did not do that. In our defense, we are still learning Open Science methods and are only now beginning the practice of developing, and publicly posting, pre-analysis plans in our research. We again note that we did not approach our replication with the goal of trying to refute AKW's findings. Rather, we attempted to be objective, independent auditors.

We note, however, that it is quite common to divide samples by geographical region. For example, growth studies will often report results for subsamples by continent (e.g., Africa, South America, Asia); or by state of economic development (e.g., OECD versus non-OECD countries; developed versus developing countries). Thus, it seemed natural for us to divide the Chinese sample into geographical regions. AKW's analysis pooled all the respective subsamples, indicating that they believed their model was applicable for firms in all regions. Separating out their sample into subsamples thus should not be a problem.

Another concern is our decision to use pre-2005 China Customs Data. AKW make a convincing case that the 2005 data is better suited for testing their theory because of the existence of licensing requirements. As they note, licensing requirements were reduced in steps

[7] "East1" is the Bohai Bay Economic Rim, including Beijing, Tianjin, Hebei, Liaoning, Shandong; "East2" is the Yangtze River Delta Economic Zone, including Shanghai, Jiangsu, Zhejiang; and "East3" is the Pearl River Delta Economic Zone, including Guangdong.

over the period 2003–2005. To the extent licensing restrictions worked against finding support for AKW's theory, the results in favor of their theory should be stronger as one moves from 2000 to 2004. However, that is not the pattern evident from Table 5. The "best" years in terms of support for Prediction #2 are 2001 and 2002. Coefficient estimates from 2003 and 2004 are wrong-signed, with the estimate from 2003 being significant at the 1-percent level.

8 Robustness check: Prediction #3

Description. The third and most important of AKW's predictions is that the share of trade carried by intermediaries is systematically related to country characteristics. As evidence in favor of their model, AKW report that the share of trade accounted by intermediaries is positively related to distance, tariffs, and number of required import documents; and negatively related to GDP and the size of the Chinese population in the destination country.

Results. The first set of additional tests for Prediction #3 repeats AKW's analysis of Table 3, this time dividing the full sample into East1, East2, East3, Central, and West subsamples. The results are reported in Panel A of Table 6A. Our benchmark specification is AKW's full model with Log Distance, Log GDP, Log Chinese Population, # of Importing Procs, MFN Tariff, and HS6 fixed effects all included in the same regression (cf. Column 4.a in Table 3). The estimated coefficients for the country characteristics are reported row-wise, with the variable predictions listed at the top of the columns. All regressions use 2005 data.

For each of the five country characteristic variables, AKW's prediction is tested by five estimates, one for each subsample. For distance and GDP, AKW's prediction is successful in four out of five, and five out of five tests, respectively. In contrast, AKW's predictions for Chinese population, number of importing procedures, and tariffs are successful in only one out of five, three out of five, and one out of five tests. This is a result we will encounter frequently in the tests ahead: strong support for AKW's Prediction #3 for distance and GDP, but weak support for Chinese population, number of importing procedures, and tariffs.

Our next set of tests focuses on different measures of the number of required importing procedures variable. AKW report that they sourced these data from the World Bank's Doing Business Report 2006, and that the variable measures "the number of procedures required for importing a container" (page 80). Our own research identified two versions of the Doing Business Report 2006, a printed book version and an online version. These were both different from each other, and different from AKW's data.

Table 7 reports summary statistics for the three versions of the number of required import documents variable. While the three versions are similar, the summary statistics reveal significant differences. This is most apparent in the pairwise comparisons reported in the last two columns. A comparison of the AKW and Book data reveal that 61.6% of the corresponding values in the two data sets are different from each other. 72.2% of the values in the AKW and Online data sets are different, and 85.4% of the values in the Book and Online versions are different. Accordingly, we investigate whether these differences affect the tests of AKW's model.

Table 6A: Robustness check of AKW's results: Prediction #3 (2005)

Sample	Distance (Expect = Pos)	GDP (Expect = Neg)	Chinese (Expect = Neg)	# Procedures (Expect = Pos)	Tariffs (Expect = Pos)	Obs.	R-sq
A) Geographical subsamples							
East1	0.038*** (0.009)	-0.012*** (0.003)	-0.002 (0.002)	0.004** (0.002)	0.011 (0.027)	69,097	0.22
East2	0.039*** (0.008)	-0.017*** (0.003)	-0.006*** (0.001)	0.005*** (0.002)	0.006 (0.027)	100,892	0.21
East3	0.019* (0.010)	-0.015*** (0.003)	-0.001 (0.002)	0.005*** (0.002)	0.054 (0.036)	68,609	0.21
Central	0.037** (0.016)	-0.011** (0.005)	-0.001 (0.003)	0.006 (0.004)	0.066 (0.059)	44,765	0.28
West	0.034*** (0.011)	-0.018*** (0.005)	0.000 (0.003)	0.001 (0.002)	0.164** (0.073)	33,835	0.32
B) Alternative data sources for number of required procedures							
Book	0.026*** (0.009)	-0.018*** (0.003)	-0.004*** (0.001)	0.001 (0.001)	0.060** (0.023)	188,120	0.17
Online	0.029*** (0.009)	-0.016*** (0.002)	-0.004*** (0.001)	0.004*** (0.001)	0.048** (0.023)	189,030	0.17
C) Alternative data sources for number of required procedures by geographical subsample							
Book/East1	0.036*** (0.009)	-0.013*** (0.003)	-0.003* (0.002)	0.002 (0.002)	0.032 (0.031)	69,097	0.22
Book/East2	0.035*** (0.008)	-0.017*** (0.003)	-0.007*** (0.001)	0.003* (0.001)	0.033 (0.030)	100,892	0.21
Book/East3	0.014 (0.010)	-0.016*** (0.004)	-0.001 (0.002)	0.002 (0.002)	0.086** (0.038)	68,609	0.21
Book/Central	0.034** (0.015)	-0.011** (0.005)	-0.001 (0.002)	0.003 (0.003)	0.092 (0.069)	44,765	0.28
Book/West	0.032*** (0.011)	-0.020*** (0.005)	0.000 (0.003)	-0.003 (0.002)	0.195** (0.079)	33,835	0.32
Online/East1	0.036*** (0.011)	-0.012*** (0.003)	-0.003 (0.002)	0.001 (0.002)	0.037 (0.033)	69,292	0.22
Online/East2	0.038***	-0.015***	-0.006***	0.005***	0.032	101,387	0.21

Sample	Distance (Expect = Pos)	GDP (Expect = Neg)	Chinese (Expect = Neg)	# Procedures (Expect = Pos)	Tariffs (Expect = Pos)	Obs.	R-sq
	(0.009)	(0.003)	(0.001)	(0.001)	(0.029)		
Online/East3	0.019* (0.010)	-0.013*** (0.003)	-0.000 (0.002)	0.005*** (0.002)	0.073* (0.038)	68,818	0.21
Online/Central	0.034** (0.015)	-0.012** (0.005)	-0.001 (0.003)	0.002 (0.003)	0.106* (0.061)	44,839	0.28
Online/West	0.033*** (0.010)	-0.017*** (0.005)	0.000 (0.003)	0.001 (0.002)	0.168** (0.070)	33,862	0.32

NOTES**:** The five rows of Panel A in the table ("East1", "East2", "East3", "Central", "West") re-estimate specification (4.a/4.b) from Table 3 using the same data, except that the full sample is divided into five mutually exclusive subsamples. The two rows of Panel B ("Book", "Online") re-estimate specification (4.a/4.b) from Table 3 using the same data except AKW's data for number of required import documents ("# of Procedures") is replaced with the printed book and online versions, respectively, of the same variable from Doing Business Report 2006. The ten rows of Panel C in the table take the full samples for the "Book" and "Online" regressions of Panel B, divide each into the five geographical subsamples from Panel A, and re-estimate specification (4.a/4.b) from Table 3. The dependent variable in each regression is the share of intermediary exports of total country HS6 exports. Regressions are estimated using OLS with cluster robust standard errors grouped by country. All regressions also include fixed effects for product type (HS6). "Pos" and "Neg" refer to positive and negative. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively. The red-shaded cells indicate that the respective estimates have the wrong sign and are statistically significant at the 5-percent level. The rose-shaded cells indicate that the respective estimates have the wrong sign, but are not statistically significant at the 5-percent level. The gray-shaded cells indicate that the respective estimates have the predicted sign, but are not statistically significant at the 5-percent level.

Table 6B: Robustness check of AKW's results: Prediction #3 (2000–2004 and Panel)

Sample	Distance (Expect = Pos)	GDP (Expect = Neg)	Chinese (Expect = Neg)	# Procedures (Expect = Pos)	Tariffs (Expect = Pos)	Obs.	R-sq
A) Different years with alternative data sources for number of required procedures							
2004/AKW	0.039*** (0.009)	-0.023*** (0.004)	-0.002 (0.002)	0.002* (0.001)	-0.007 (0.007)	106,033	0.23
2003/AKW	0.047*** (0.006)	-0.021*** (0.004)	-0.006*** (0.002)	0.005*** (0.001)	0.008 (0.009)	99,608	0.24
2002/AKW	0.046*** (0.008)	-0.020*** (0.004)	-0.004** (0.002)	0.005*** (0.002)	-0.006*** (0.002)	96,785	0.24
2001/AKW	0.046*** (0.007)	-0.017*** (0.003)	-0.005*** (0.002)	0.006*** (0.002)	-0.008 (0.009)	91,491	0.24
2000/AKW	0.042*** (0.008)	-0.022*** (0.004)	-0.008*** (0.002)	0.006*** (0.001)	-0.004 (0.012)	82,396	0.26
2004/Book	0.037*** (0.009)	-0.025*** (0.004)	-0.002 (0.002)	-0.000 (0.001)	-0.004 (0.007)	106,033	0.23
2003/Book	0.045*** (0.007)	-0.022*** (0.004)	-0.005** (0.002)	0.002* (0.001)	0.014 (0.010)	99,608	0.24
2002/Book	0.043*** (0.009)	-0.023*** (0.004)	-0.004 (0.002)	0.002 (0.001)	-0.004* (0.002)	96,785	0.23
2001/Book	0.043*** (0.008)	-0.020*** (0.004)	-0.004* (0.002)	0.001 (0.002)	0.009 (0.014)	91,491	0.24
2000/Book	0.040*** (0.008)	-0.024*** (0.005)	-0.008*** (0.002)	0.003* (0.001)	0.006 (0.013)	82,396	0.26
2004/Online	0.040*** (0.009)	-0.022*** (0.003)	-0.002 (0.002)	0.003** (0.001)	-0.005 (0.007)	107,110	0.23
2003/Online	0.047*** (0.008)	-0.020*** (0.004)	-0.005** (0.002)	0.004** (0.002)	0.018 (0.011)	99,677	0.24
2002/Online	0.045*** (0.010)	-0.020*** (0.004)	-0.003 (0.002)	0.005*** (0.002)	-0.005** (0.002)	96,907	0.23
2001/Online	0.045*** (0.008)	-0.017*** (0.004)	-0.004* (0.002)	0.005*** (0.002)	0.010 (0.014)	92,197	0.24
2000/Online	0.041***	-0.022***	-0.007***	0.005***	0.009	82,443	0.26

Sample	Distance (Expect = Pos)	GDP (Expect = Neg)	Chinese (Expect = Neg)	# Procedures (Expect = Pos)	Tariffs (Expect = Pos)	Obs.	R-sq
	(0.009)	(0.004)	(0.002)	(0.001)	(0.012)		
B) Panel fixed effects with alternative data sources for number of required procedures							
Panel/ AKW (With Year FEs)	0.040*** (0.004)	-0.019*** (0.002)	-0.005*** (0.001)	0.005*** (0.001)	-0.002 (0.003)	664,433	0.18
Panel/Book (With Year FEs)	0.037*** (0.004)	-0.021*** (0.002)	-0.004*** (0.001)	0.002*** (0.001)	0.002 (0.005)	664,433	0.17
Panel/Online (With Year FEs)	0.039*** (0.004)	-0.019*** (0.001)	-0.004*** (0.001)	0.004*** (0.001)	0.002 (0.005)	667,364	0.17

NOTES**:** The fifteen rows of Panel A re-estimate specification (4.a/4.b) from Table 3 using earlier years (2000-2004) of the China Customs Data. It does this for each of the three versions of the "# of Procedures" variable ("AKW", "Book", "Online"). Note that the same values that were used for the variables "Chinese" and "# of Procedures" in the 2005 regressions are also used for the earlier years because data for 2000-2004 for these variables are unavailable. The dependent variable in each regression is the share of intermediary exports of total country HS6 exports. Regressions are estimated using OLS with cluster robust standard errors grouped by country. All regressions also include fixed effects for product type (HS6). "Pos" and "Neg" refer to positive and negative. *, **, and *** indicate statistical significance at the 10-, 5-, and 1-percent level, respectively. The red-shaded cells indicate that the respective estimates have the wrong sign and are statistically significant at the 5-percent level. The rose-shaded cells indicate that the respective estimates have the wrong sign, but are not statistically significant at the 5-percent level. The gray-shaded cells indicate that the respective estimates have the predicted sign, but are not statistically significant at the 5-percent level.

Table 7: Summary statistics for three data sources for the number of required import documents variable

Source	Observations	Mean	Median	Min	Max	% Different: AKW	% Different: Book
AKW	151	9.10	9	3	19	----	----
Book	151	10.77	11	3	19	61.6	----
Online	171	8.08	7	2	21	72.2	85.4

NOTES: This table compares the three versions of the variable "# of Procedures" available from AKW, the printed book version of the World Bank's Doing Business Report 2006, and the online version of the World Bank's Doing Business Report 2006. The last two columns compare the percent of total number of observations that are different from each pair of versions (AKW/Book, AKW/Online, Book/Online). Percentages are calculated from the total number of paired observations in which both versions report non-missing values.

Panel B of Table 6A reports the results of substituting the book and online versions of the number of required import documents for AKW's version in the 2005, full sample regressions. Focusing first on the number of procedures variable, the estimated coefficient for the printed book version is correctly signed, but insignificant. The corresponding estimate for the online version is virtually identical to the estimate using AKW's data (see Column 4.b in Table 3). The tests for the other country characteristics are qualitatively unaffected by using alternative versions of the required documents variables, with all estimates confirming AKW's model.

Panel C repeats the preceding analysis, except that it divides the full sample into regional subsamples. This produces a total of ten tests for each variable. As we have consistently seen previously, the estimates for distance and GDP continue to strongly support AKW's model. In contrast, the estimates for Chinese population, number of procedures, and tariffs are substantially weaker, with only two out of ten, two out of ten, and three out of ten of the estimates providing evidence in favor of AKW's predictions, where confirmatory evidence is defined as a correctly signed estimate that is significant at the 5 percent level.

Our last set of robustness tests exploits the fact that we have Chinese Customs Data for the years 2000 to 2004. Unfortunately, we do not have time-varying data for Chinese population or number of required import documents. The variable measuring Chinese population is obtained from the Ohio University Shao Centre. This variable is collected in different regions in different time periods. For many countries, the data were collected prior to 2005. Thus, they are arguably as applicable, or more, for these earlier years. Relatedly, our only data for number of required import documents comes from the 2006 edition of Doing Business Reports. This reports information current as of January 2005. As a result, we must assume that these data are also valid for 2000 to 2004.

With these caveats in mind, Table 6B/Panel A reports the results of estimating the bench-mark specification for the years 2000-2004, using the three different measures of number of import documents (AKW, Book, and Online). The result is a total of 15 tests of each variable. Once again, the coefficients for distance and GDP show strong support for AKW's theory, with every regression producing coefficients that are correctly signed and statistically significant.

With respect to the Chinese population coefficient, the results continue to be favorable to AKW's model when using their measure of required documents. The respective coefficient is correctly signed and statistically significant in four of the five regressions (2000–2003). When alternative versions of the required documents are substituted for AKW's version, this falls to

two out of five (2000 and 2003 for both the Book and Online versions). For the required documents coefficient, both AKW's and the Online data provide strong support for AKW's model for the years 2000–2004, producing correctly-signed, significant coefficients in four out of five and five out of five cases. In contrast, when the data for number of required import documents come from the printed book version, none of the five estimates provides supporting evidence because all are statistically insignificant. Lastly, the results for the tariff variable are easy to summarize: None of the 15 estimates provide support for AKW's model. In fact, two of the 15 estimates are wrong-signed and significant (2002/AKW and 2002/Online).

Table 6B/Panel B provides one final set of robustness tests of AKW's Prediction #3. Here we pool data over the years 2000–2005 and estimate a panel, two-way fixed effects model using the three different measures of number of required import procedures. The associated results produce strong confirmatory evidence for AKW's predictions for four of the five variables: distance, GDP, Chinese population, and number of import procedures. All of the estimated coefficients are correctly signed and statistically significant. The exception are the estimates for tariffs, where all three coefficients are statistically insignificant.

With these caveats in mind, Table 6B/Panel A reports the results of estimating the benchmark specification for the years 2000–2004, using the three different measures of number of import documents (AKW, Book, and Online). The result is a total of 15 tests of each variable. Once again, the coefficients for distance and GDP show strong support for AKW's theory, with every regression producing coefficients that are correctly signed and statistically significant.

With respect to the Chinese population coefficient, the results continue to be favorable to AKW's model when using their measure of required documents. The respective coefficient is correctly signed and statistically significant in four of the five regressions (2000–2003). When alternative versions of the required documents are substituted for AKW's version, this falls to two out of five (2000 and 2003 for both the Book and Online versions). For the required documents coefficient, both AKW's and the Online data provide strong support for AKW's model for the years 2000–2004, producing correctly-signed, significant coefficients in four out of five and five out of five cases. In contrast, when the data for number of required import documents come from the printed book version, none of the five estimates provides supporting evidence because all are statistically insignificant. Lastly, the results for the tariff variable are easy to summarize: None of the 15 estimates provide support for AKW's model. In fact, two of the 15 estimates are wrong-signed and significant (2002/AKW and 2002/Online).

Table 6B/Panel B provides one final set of robustness tests of AKW's Prediction #3. Here we pool data over the years 2000–2005 and estimate a panel, two-way fixed effects model using the three different measures of number of required import procedures. The associated results produce strong confirmatory evidence for AKW's predictions for four of the five variables: distance, GDP, Chinese population, and number of import procedures. All of the estimated coefficients are correctly signed and statistically significant. The exception are the estimates for tariffs, where all three coefficients are statistically insignificant.

Summary. The record with respect to distance and GDP provides very strong evidence of AKW's Prediction #3. For GDP, every estimated coefficient in Tables 6A and 6B lines up with AKW's prediction (35 out of 35). For distance, all of the estimated coefficients are correctly signed, and 32 out of 35 are statistically significant. The overall record in Tables 6A and 6B with respect to the other three country characteristics is generally not supportive: In

approximately half or more of the tests, the estimated coefficients do not support AKW's prediction.

Qualifications. As noted above, we do not have time-varying data for Chinese population and number of required import documents for the years 2000–2004. This potentially introduces measurement error in these regressions that could bias the coefficients towards zero. Accordingly, one could argue that the associated tests should receive less weight.

If we omit the results from Table 6B, rather than 16/35, 18/35, and 6/35 of the estimates supporting the predictions for Chinese population, number of required documents, and tariffs, the numbers are 5/17, 6/17, and 6/17. Further, if one believes that among these, the most reliable results are those that use the Online measure of import procedures, as the reviewers recommend, the corresponding numbers are 2/6, 3/6, and 2/6. The only supporting results are those from the panel fixed effects analysis, where the predictions for all five country characteristics are confirmed. However, the online comments from the original authors discourage the use of the pre-2005 data.

9 Overall assessment of the robustness checks

Table 8 pulls together all the test results for AKW's three predictions, both (i) reproduction and (ii) re-analysis and extension. Our reproduction of AKW confirms AKW's conclusion that the empirical analysis provides strong support for their model. 6/9 test results support Prediction #1, 3/3 test results support Prediction #2, and 13/14 test results support Prediction #3. However, when we re-analyze and extend the analysis, dividing the data by geographical regions, using different versions for number of required import documents, and extending the data to additional years, the evidence becomes much weaker.

With respect to Prediction #1, less than half of the predictions (4/9) are supported by the data. Likewise for Prediction #2: only three out of eight predictions are supported by the cross-sectional estimates. The panel fixed effects estimate also does not provide support.

AKW identify Prediction #3 as their "central prediction". Here the results are mixed. Overall, 107 out of 175 predictions support the theory. However, these mask an important difference within the five country characteristics examined by AKW. While the test results for distance and GDP strongly support AKW's theory, those for Chinese population, number of import procedures, and tariffs, do not. Even when we restrict the results to tests where the underlying data are judged to be more reliable (2005), and we use the preferred measure for the number of import procedures variable (Online), the numbers of successful predictions are never more than half.

Table 8: Summary of results from tests of AKW's model

PREDICTION	**REPRODUCTION**		**RE-ANALYSIS AND EXTENSION**					
Prediction #1	Table 1:	**6/9**	Table 4: **4/9**					
Prediction #2	Table 2:	**3/3**	Table 5 (All except Panel FE):	**3/8**		Table 5 (Panel FE):		**0/1**
Prediction #3 – Total	Table 3:	**13/14**	Tables 6A+6B (All):	**107/175**	Table 6A:	**48/85**	Table 6A (Online):	**4/5**
Prediction #3 – Distance	Table 3:	**4/4**	Tables 6A+6B (All):	**32/35**	Table 6A:	**14/17**	Table 6A (Online):	**5/6**
Prediction #3 – GDP	Table 3:	**4/4**	Tables 6A+6B (All):	**35/35**	Table 6A:	**17/17**	Table 6A (Online):	**6/6**
Prediction #3 – Chinese	Table 3:	**3/3**	Tables 6A+6B (All):	**16/35**	Table 6A:	**5/17**	Table 6A (Online):	**2/6**
Prediction #3 – # Procedures	Table 3:	**2/2**	Tables 6A+6B (All):	**18/35**	Table 6A:	**6/17**	Table 6A (Online):	**3/6**
Prediction #3 – Tariff	Table 3:	**0/1**	Tables 6A+6B (All):	**6/35**	Table 6A:	**6/17**	Table 6A (Online):	**2/6**

NOTES**:** The values in the table report the number of tests that support the respective prediction over the total number of tests. The results in the "Reproduction" column collect the test results from Tables 1 to 3. The results in the "Re-analysis and Extension" column collect the results from Tables 4 to 6A/6B. A test was judged to support AKW's predictions if it had the predicted sign and was statistically significant. A coefficient that had the predicted sign but was statistically insignificant was interpreted as not providing statistical support for AKW's model. "Prediction #3 – Total" combines the results from predictions for the variables Distance, GDP, Chinese, # of Procedures, and Tariff. "Prediction #3 – Distance" to "Prediction #3 – Tariff" break out the results for testing the predictions of the individual variables

10 Conclusion

Mayo (2018) argues that one reason science has been vulnerable to a reproducibility crisis is because theories are weakly tested. Accordingly, she advocates for "severe testing". Only when a theory has successfully survived an appropriate number of tests should that theory be viewed as credible.

We apply this approach in our replication of Ahn, Khandelwal, and Wei´s (2011) model of intermediate trade. AKW has been influential in the literature because it provides both a theoretical framework for explaining the existence of trade intermediaries, and an accompanying empirical confirmation of the model's predictions. We employ a two-pronged approach to re-examining their empirical analysis. First, we reproduce the main results from their empirical analysis, applying the same empirical procedures to (mostly) independently sourced data. Our reproduction results strongly confirm AKW's conclusions.

We then undertake a series of re-analyses and extensions, sometimes re-working the reproduction data, other times turning to alternative data sources or extending the time frames of their analysis. When we do that, we find that empirical support for the AKW model of intermediated trade is greatly diminished.

For example, when we extend AKW's use of Enterprise Survey Data to new data from 2012, we estimate a U-shaped relationship between productivity and indirect exports, rather than the inverted U-shaped relationship predicted by their theory (Prediction #1). Another example arises when we divide the full dataset according into five separate geographical regions. We only find confirming evidence for Prediction #2 in the Pearl River Delta Economic Zone of China ("East3"). The results from the other regions are either contradictory (wrong-signed and significant) or insignificant. While our analyses find some support for Prediction #3, they only find it for two of the five country characteristics tested by AKW.

In summary, while our additional tests produce some successes for the AKW model, a holistic assessment leads to the conclusion that the data do not generally support their theory. That being said, the results reported here should not be interpreted as solely negative. They point to possible avenues for future research. A robust result in our analysis is that the share of exports through intermediaries is positively correlated with distance to trading partner, and negatively correlated with the size of the trading partner's market. This highlights the need for a theoretical explanation for why these two country characteristics in particular should be associated with intermediated trade. We hope that this research stimulates efforts in this direction.

Acknowledgments: We gratefully acknowledge assistance from Ahn, Khandelwal, and Wei. They provided their programming code, patiently answered our queries about their data, and shared their tariff and import procedures data with us. Remaining errors are our own.

References

Ahn, J., Khandelwal, A. K., and Wei, S.-J. (2011). The role of intermediaries in facilitating trade. *Journal of International Economics*, 84(1), 73–85. https://doi.org/10.1016/j.jinteco.2010.12.003

Akerman, A., 2018. A theory on the role of wholesalers in international trade based on economies of scope. *Canadian Journal of Economics/Revue canadienne d'économique*, 51(1), 156–185. https://doi.org/10.1111/caje.12319

Antràs, P., and Costinot, A. (2011). Intermediated trade. *The Quarterly Journal of Economics*, 126(3), 1319–1374. https://scholar.harvard.edu/antras/publications/intermediated-trade

Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2010). Wholesalers and retailers in US trade. *American Economic Review*, 100(2), 408–413. https://www.aeaweb.org/articles?id=10.1257/aer.100.2.408

Blum, B.S., Claro, S. and Horstmann, I.J., 2018. Trade costs and the role of international trade intermediaries. In *Handbook of International Trade and Transportation*. Edward Elgar Publishing. https://doi.org/10.4337/9781785366154.00018

Chan, J.M., 2018. Financial frictions and trade intermediation: Theory and Evidence. *European Economic Review*, https://doi.org/10.1016/j.euroecorev.2018.04.002.

Dasgupta, K. and Mondria, J. (2018). Quality uncertainty and intermediation in international trade, *European Economic Review*, 104, 68–91. https://doi.org/10.1016/j.euroecorev.2018.02.002

Feenstra, R. C., and Hanson, G. H. (2004). Intermediaries in Entrepot Trade: Hong Kong Re-Exports of Chinese Goods. *Journal of Economics and Management Strategy*, 13(1), 3–35. https://doi.org/10.1111/j.1430-9134.2004.00002.x

Mayo, D. G. (2018). *Statistical inference as severe testing*. Cambridge: Cambridge University Press.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6), 1695–1725. https://doi.org/10.1111/1468-0262.00467

Poncet, S. and Xu, M., 2018. Quality screening and trade intermediaries: Evidence from China. *Review of International Economics*, 26(1), 223–256. https://doi.org/10.1111/roie.12324

Qu, Z., Raff, H., and Schmitt, N. (2014). Inventory and International Trade Volatility: The Role of Export and Import Wholesalers. Working paper, University of Kiel. https://www.etsg.org/ETSG2014/Papers/191.pdf

Rossman, M. L. (1984). Export trading company legislation: US response to Japanese foreign market penetration. *Journal of Small Business Management* (pre-1986), 22(000004), 62.

Sheveleva, Y., and Krishna, K. (2017). Wheat or strawberries? Intermediated trade with limited contracting. *American Economic Journal: Microeconomics*, 9(3), 28–62. https://www.aeaweb.org/articles?id=10.1257/mic.20140189

Spulber, D.F. (1996). Market microstructure and intermediation. *Journal of Economic Perspectives*, 10(3), 135–152. https://www.jstor.org/stable/2138524

Tang, H. and Zhang, Y. Quality differentiation and trade intermediation. Centro Studi Luca d'Agliano Development Studies Working Paper No. 340. Available at SSRN: https://ssrn.com/abstract=2214960 or http://dx.doi.org/10.2139/ssrn.2214960 .

Appendix

Description of variables and data sources used in the reproduction replication of AKW

Variable	***Definition***	***Source***
Direct export share	Percentage of firm's sales exported directly	World Bank's Enterprise Survey Data
Indirect export share	Percentage of firm's sales exported indirectly (through a distributor)	World Bank's Enterprise Survey Data
Sales	Total sales 1 year ago	World Bank's Enterprise Survey Data
Employment	Average number of permanent workers plus temporary workers 1 year ago	World Bank's Enterprise Survey Data
Labor productivity	$\frac{sales}{employment}$	World Bank's Enterprise Survey Data
Unit value	$\frac{value\ of\ sales}{quantity\ sold}$	China Customs Data
Intermediary	= 1 if the firm is an intermediary and = 0 otherwise	China Customs Data
Firm size	Firm's total export value	China Customs Data
Intermediary export share	Share of intermediary exports of total country-HS6 exports	China Customs Data
Distance	Air distance in nautical miles between the trading country pairs	https://www.timeanddate.com/worldclock/distance.html
GDP	Gross domestic product	World Bank's World Development Indicators
Chinese population	Size of the ethnic Chinese population per country	Ohio University Shao Centre
Importing procedures	Number of documents required for import in China	World Bank's Doing Business Report (*supplied by AKW*)
MFN tariff	Most Favored Nation duty rate treatment	World Bank WITS (*supplied by AKW*)

Disclosure Statement

The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.