Germs, Roads and Trade: Theory and Evidence on the Value of Diversification in Global Sourcing

Hanwei Huang *

This version: December, 2017

[please click here for the latest version]

Abstract

This paper studies the value of diversification in global sourcing in improving firms' resilience to supply chain disruptions. I build a model in which firms select into importing from countries and via customs with different efficiencies, taking into account domestic and international trade costs. The model predicts that firms which are more geographically diversified in sourcing are more resilient to supply chain disruptions. Reductions in trade costs induce firms to further diversify their sourcing strategies. I then exploit the 2003 SARS epidemic as a natural experiment to examine the resilience of Chinese manufacturing importers. Firm imports fell by 7.9% on average when the trade route was hit by SARS, but as much as 52% for firms without any diversification. The disruptions led to smaller increases in marginal cost for firms with more trade routes for imports. The epidemic reduced total Chinese manufacturing outputs by about 0.7% at its peak. Connectivity to roads increased firms' resilience to the epidemic by facilitating input diversification.

Key Words: Global sourcing, Diversification, Resilience, Firm heterogeneity, SARS

*London School of Economics, E-mail: h.huang7@lse.ac.uk. I am indebted to Robin Burgess, Gianmarco Ottaviano, Thomas Sampson, and Catherine Thomas for their advice and constant support. I would like to thank Jiandong Ju and Xinding Yu for the help with data access. For insightful discussions, I would like to thank Karun Adusumilli, Andy Bernard, Giuseppe Berlingieri, Marcus Biermann, Emily Blanchard, Holger Breinlich, Lorenzo Caliendo, Paola Conconi, Fabrice Defever, Jonathan Dingel, Dave Donaldson, Peter Egger, Chao He, Giammario Impullitti, Kalina Manova, John Morrow, Dennis Novy, Emanuel Ornelas, Frank Pisch, Veronica Rappoport, Zheng Michael Song, Alan Spearot, Daniel Sturm, Gonzague Vannoorenberghe, Miaojie Yu, Shengxing Zhang, and Yuan Zi. I would also like to thank participants of the LSE trade seminars, the CEP annual conference, ETSG 2017, and UCLouvain for their useful comments. All the errors are my own responsibility. A previous version of this paper was circulated as: "Germs, Roads and Trade: Resilience of Globally Sourcing Firms in the Face of the SARS Epidemic".

1 Introduction

Global sourcing has allowed firms to find the best input in a global market but also exposed them to foreign shocks. For example, the 2011 Tōhoku earthquake in Japan caused severe disruptions to affiliates of Japanese multinationals in the US (Boehm, Flaaen and Pandalai-Nayar, 2015). Despite the fact that firms are increasing the priority of supply chain management and the conventional wisdom suggesting firms to diversify, there is little rigorous evidence on how diversification in global sourcing shapes the impact of supply chain disruptions on firms and the extent to which infrastructure affects the size of the impact.[1]

In this paper I study the value of diversification in global sourcing for Chinese manufacturing importers by exploiting the 2003 SARS epidemic as a natural experiment. I show both in theory and empirics that geographical diversification is crucial in building a resilient supply chain.[2] By doing this I make the following three theoretical and empirical contributions. First, I find that high productivity firms are more *geographically* diversified in input sourcing than low productivity firms. Second, I find that sourcing diversifications make firms more resilient to adverse shocks on sourcing routes under input complementarity. Finally, I find that connectivity to transportation networks increases sourcing diversification by inducing firms to source via more trade routes, which helps to dampen the negative impact of adverse shocks.

The 2003 SARS epidemic provides the empirical setting for my investigation on supply chain disruptions. Unlike the recent outbreaks of Ebola and Zika, Severe Acute Respiratory Syndrome (SARS) was an unknown disease when it first struck southern China in late 2002. It rapidly hit several other countries/regions, and reached its peak in the second quarter of 2003. The epidemic ended in July 2003, after affecting more than 8,000 and taking away the life of 774 people. The rapid spread, coupled with scant information disclosed by the Chinese government, shocked the global community. Major trading partners of mainland China such as Canada, Hong Kong, Taiwan and Singapore, and trade hubs in China such as Beijing and Guangdong were severely affected. Given its deadliness and infectiousness, governments took stringent measures

[1] More than 90% of the firms surveyed by the World Economic Forum (2012) indicated that supply chain and transport risk management had become a greater priority for them. A Financial Times article (2014) advocates that diversification is still at the heart of supply chain management. However, management and operation scientists mostly rely on simulations to evaluate supply chain disruptions and have problems estimating model parameters according to the review by Snyder et al. (2016).

[2] A firm is defined to be more resilient if pass-throughs of adverse shocks on trade routes to firm outcomes (such as marginal costs or revenues) are smaller. Ponomarov and Holcomb (2009) survey other notions of resilience.

to combat SARS, including travel bans,[3] vessel controls at ports,[4] health check-points on roads, and quarantine of SARS contacts, which inevitably disrupted trade. For example, the number of visitors to the 2003 spring session of Canton Fair, the largest trade fair in China, dropped by 81%, and the total business turnover dropped by 74% year-on-year.[5]

To guide the empirical analysis, I build a model in which firms with heterogeneous productivity source inputs from different places to assemble final goods. In making the sourcing decisions, a firm first chooses the sourcing routes. Conditional on its established routes, the firm then chooses imports across this set of routes. The model has the following key testable predictions. First, given the assumption that adding new sourcing routes incurs fixed costs, only high productivity firms can afford to source via more routes under input complementarity. I further show that they are more diversified in sourcing than low productivity firms as measured by the Herfindahl-Hirschman Index which sums over the squares of input expenditure share of each route. Second, more diversified firms are more resilient to adverse shocks under input complementarity. I show that the pass-through of an adverse shock on sourcing routes to marginal cost is proportional to the input expenditure share of the route hit by the shock, which tends to be smaller for more diversified firms. The rise in marginal costs drives down input demands and imports under input complementarity, *vice versa* if inputs are substitutable. Such a feedback effect from marginal cost to imports is smaller for a more diversified firm. Therefore, estimating the size of the pass-through to imports not only identifies the extent to which inputs are complementary but also tests the diversification channel. Finally, the model predicts that reduction in trade costs induces firms to further diversify by sourcing from more places.

I test the model prediction on diversification and resilience by estimating the response of Chinese manufacturing importers to SARS using matched customs and firm data from 2000-2006. The data allow me to identify the date, the location of the importer, the Chinese entry customs, and the origin for each transaction. To capture the spatial and time variations of the epidemic, I construct a treatment variable which measures the exposure of Chinese importers to SARS by sourcing route. A route is defined as affected if the origin or the entry customs district was on the WHO's list of areas with local SARS outbreaks. Since the model predicts that the

[3] The World Health Organization (WHO) issued rare travel advice warning travellers against visiting regions with local outbreaks (Heymann, Mackenzie and Peiris, 2013).

[4] The WHO also provided guidelines to port authorities if cruise vessels had suspected cases on board. The number of vessels arriving in Hong Kong dropped by about 5% in the first half of 2003. A Malaysian chemical cargo vessel heading to Guangzhou was held in quarantine for 10 days when the crew members started developing SARS-like symptoms. More than two months elapsed before the sick crew members were given the all-clear.

[5] Source: Historical statistics of the Canton Fair.

pass-through of a trade cost shock into imports depends on the pre-shock input expenditure share of the route got hit, I include an interaction term between the treatment variable and the average input expenditure share by route before SARS to capture such heterogeneous treatment effect. Because SARS struck some parts of China which might affect demands and productions directly, I include firm-time fixed effects to capture firm-level time-varying demand and productivity shocks, and rely on variations within firms across sourcing routes to identify the effect of SARS. The baseline estimates imply that the average pass-through of the SARS shock to imports was about -7.9%. Crucially, the impact increased with the pre-SARS input expenditure share which implies that inputs were complementary and diversification brought resilience. For a firm solely relied on the route hit by SARS, its import would drop by as much as 52%.

More diversified firms saw smaller impacts on their route-specific imports, but the overall impact might not be smaller if a larger number of their sourcing routes were affected. To see if that was the case, I use the model to account the effect of SARS on other firm level outcomes. Despite the fact that I do not have the data to do a full-fledged structural estimation, I show that we can gauge the effect on firm marginal costs and outputs using a sufficient statistic approach. The idea is to combine the "hat algebra" approach (Jones, 1965; Dekle, Eaton, and Kortum, 2007) and the technique from Feenstra (1994).[6] Using this new method, I find that the marginal cost of firms whose imports were hit by SARS increased by about 0.7% on average. The rise in marginal costs tended to be smaller for firms with more sourcing routes. Conversely, if pass-throughs were homogeneous, firms with more sourcing routes would be more heavily affected. Aggregating across firms, total output decreased by about 0.7% at the peak of SARS.[7]

The model predicts that high productivity firms are more geographically diversified which I confirm in the data. Conditional on productivities, the model also predicts that firms' sourcing strategies expand weakly if trade costs decline. Therefore, improvements in infrastructure which reduce trade costs induce firms to become more diversified. This might make them more resilient to adverse shocks given the finding that diversification brings resilience. To test these model implications, I exploit the expansion of Chinese highway and railway networks from 2000-2006 and examine whether firms further diversify their sourcing strategy or not after connecting to railways or highways. Indeed, I find that firms located in regions connected to highways started

[6]Feenstra (1994) found that we can estimate changes in the price index even if there are new or disappearing varieties as long as there are varieties which are available both before and after. Similarly, I estimate changes in firms' marginal cost relying on overlapping trade routes prior and post the shock.

[7]It is about two thirds of the GDP loss estimated by Lee and McKibbin (2004) using a CGE model. I do not consider input-output linkages which could amplify the effect as in Carvalho et al. (2016).

to source via more trade routes, but connectivity to railways only had significant effects on the intensive margin. To deal with the potential endogeneity of highway or railway placements, I follow the "inconsequential unit approach" to exclude regions located on nodes of the transportation network and focus on the periphery regions (Chandra and Thompson, 2000). The effect remains robust and significant. Finally, I provide evidence that connectivity to railways dampened the negative impact of SARS on imports by about 6% for firms in the periphery regions while the effect of highway connection was insignificant.

I conduct various robustness checks on the baseline result. First, to deal with concerns over omitting export demand shocks, I extend the benchmark model by allowing firms to export and derive a new structural equation incorporating export demand shocks. Guided by the extended model, I construct controls for export demand shocks. The estimated effect of export demand shocks turns out to be small and insignificant. Second, I need to ensure that the SARS shock was as good as random to firms in order to estimate its effect consistently. To test this assumption, I employ a Difference-In-Difference strategy to show that the growth trends of the never-treated and eventually-treated imports were similar before SARS. Third, to deal with concerns on the peculiar feature of processing trade and its prominence in Chinese imports, I estimate the response of importers doing Processing with Inputs (PI) and Pure Assembly (PA), separately. PA firms do not decide where to source or own the imported inputs but must have written contracts approved by the customs authority in advance (Feenstra and Hanson, 2005). There is little scope for them to adjust sourcing in the face of SARS. Indeed, I find no significant treatment effect or differential treatment effect for PA firms, while the diversification channel still works for PI firms. Finally, I examine the possibility of alternative mechanisms cushioning firms from negative shocks. I construct variables to measure firms' inventories, access to finance and liquidity, and include them with the diversification channel. The diversification channel remains robust but these alternative mechanisms are insignificant. To deal with multi-plant firms diversifying productions in multiple locations, I focus on firms importing/exporting in a single location and find the diversification mechanism remains significant for them.[8]

Related Literature

My paper is related to several strands of the literature. It first contributes to studies on trade in intermediate inputs and global value chains. There is a large body of literature studying the

[8]Neither the firm survey nor the customs data report the number of plants that a firm has. The customs data report the origin for each export and destination for each import transaction. I use a proxy which counts the distinct number of Chinese locations (approximately at county level) associated with each firm.

productivity and welfare gains from sourcing foreign intermediate inputs (Hummels et al., 2001; Goldberg et al., 2010; Gopinath and Neiman, 2013; Halpern et al., 2015; Yu, 2015; Blaum et al., 2016). This paper builds on the work by Antràs, Fort, and Tintelnot (2017, hereafter AFT) and highlights another benefit of global sourcing, namely allowing firms to diversify their sourcing strategies and increase their resilience to adverse shocks.[9] While firm heterogeneity has been shown to affect firms' organizational form (Antràs and Helpman, 2004) and the productivity gains of sourcing (Blaum et al., 2016), I show that it also shapes firms' resilience to supply chain disruptions and output volatility.[10]

My study is also related to the literature on diversification in trade. The mechanism of my model is similar to the "technological diversification" channel in Koren and Tenreyro (2013). They show that it can explain the differential country-level output volatilities in a close-economy model with endogenous growth. I show how it can generate resilience to supply chain disruptions and heterogeneous firm-level volatility in open economies.[11] Whereas only the extensive margin is active in their model, I look at diversification in both the intensive and extensive margins. Allen and Atkin (2016) investigate how the expansion of Indian highways has shaped farmers' revenue volatility and crop allocations through the lens of a model with risk-averse agents, but technological diversification is achieved by risk-neutral agents in my model. Relatedly, using models with risk-averse agents, Fillat and Garetto (2015), Esposito (2016), and Kramarz et al. (2016) examine demand diversifications for multinationals and exporters. I focus on the sourcing diversification mechanism and test it using a natural experiment.

The paper also contributes to the lively literature evaluating the impact of natural disasters or epidemics on economic activity (Young 2005; Hsiang and Gina 2014; Boehm et al. 2015; Barrot and Sauvagnat 2016; Carvalho et al. 2016). Similar to Boehm et al. (2015), Barrot and Sauvagnat (2016), and Carvalho et al., (2016), I also study how shocks affect the rest of the economy or other economies through the input channel. The key difference is that I focus on firms' heterogeneous response and how diversification can serve as a mechanism to mitigate negative shocks. While the detrimental effect of Ebola on trade has been noted (FAO, 2016;

[9] Similar to AFT, Bernard et al. (2015), Blaum et al. (2016), and Furusawa et al. (2015) also study firms' extensive margin choice in sourcing but with different focuses from mine.

[10] AFT investigates the heterogeneous response of US firms to a large long-term shock which increases the potential of China in supplying intermediate inputs. Berman et al. (2012) and Amiti et al. (2014) both study how firm heterogeneity matters for exporters' response to exchange rate shocks. I focus on importers' heterogeneous response to a negative temporal real shock.

[11] Di Giovanni and Levchenko (2012), and Caselli et al. (2014) study country-level volatility in open economies. Vannoorenberghe et al. (2016), Kurz and Senses (2016) also found firm-level volatilities are related to exporting and importing.

World Bank, 2016), there is little concrete evidence of this effect. This paper is thus the first to evaluate the impact of an epidemic on trade in intermediate inputs.

Finally, the paper is related to studies on infrastructure and trade. While most of the literature focuses on how infrastructure reduces trade costs and brings productivity or welfare gains,[12] my study highlights an additional benefit of better infrastructure, that is, allowing firms to diversify sourcing and increase their resilience to shocks. Similar effects of infrastructure are also featured in Burgess and Donaldson (2010, 2012) who find that the arrival of railways in India reduced the damage of weather shocks on local economies.

The remainder of the paper is organized as follows. Section 2 presents the motivating evidence. Section 3 sets up the model and develops its main predictions. Section 4 studies the resilience of firms to SARS. Section 5 accounts for the effect on marginal costs and revenues. Section 6 examines the effect of roads on diversification and resilience. Section 7 concludes.

2 Motivating Evidence

This section establishes three new stylized facts on global sourcing which motivates the theoretical model in the next section. I use two datasets to generate these facts. The first is the Chinese Annual Industry Survey (CAIS) for year 1999-2007. It covers all state owned enterprises and other firms with sales above 5 million Chinese Yuan (around US$60,000). It provides firms' financial statements, name, address, phone number, and post code etc.. The other data that I use are the Chinese Customs data for year 2000-2006 which cover all Chinese import and export transactions. For each transaction, the data record the value, quantity, origin, destination, the Chinese customs district for clearance, and information about the Chinese import/export entity. There is no common identifier between these two datasets. I match them using firm name, post code, and phone number.[13] Because my focus is the production of goods, I limit the sample to manufacturing firms. Firms with fewer than 8 employees are excluded since they operate under different legal requirements. I also exclude firms with negative outputs or fixed assets. The matched sample represents about 38% of all Chinese imports in 2000 and 46% of those in 2006.

[12]Recent contributions include Donaldson (forthcoming), Allen and Arkolakis (2014), Fajgelbaum and Redding (2014), Atkin and Donaldson (2014), Bernard et al. (2015), and Baum-Snow et al. (2016).

[13]This matching method has been used in various papers including Yu (2015), and Manova and Yu (2016).

2.1 Output Volatility and Sourcing Diversification

Since the customs data record the origin, destination, and customs district, I can track the geographical trajectory of each transaction.[14] For example, a firm from Beijing can import from Japan via the Shanghai or the Tianjin customs district. In total, China was divided into 41 provincial level customs districts during the sample period.[15] The majority of these customs districts overlap the provincial borders as shown on the map (see Figure A2 in the Appendix).[16] A full list of the custom districts is given in Table A10.

The combination of a sourcing origin and a customs district forms a geographically distinct route for sourcing. Using this information, I first identify the set of sourcing routes used by each Chinese importer. I then measure sourcing diversification for each firm using the Herfindahl-Hirschman Index (HHI) which sums over the squares of input expenditure share of all routes, while the input expenditure share is measured by the share of route-specific inputs in total inputs. Since domestic sourcing is not observed in my data, HHI is assigned as one for non-importers. At the same time, CAIS allows me to compute the volatility of outputs for firms. Following Koren and Tenreyro (2013), I define output volatility as the variance of (real) sales growth rate during the period 1999 to 2007.[17] Since this is a relatively short time series, I also use the customs data to generate a relatively long time series of firms' quarterly exports and compute the volatility of exports for exporters from 2000-2006.[18] I then examine how sales and exports volatility are associated with firms' sourcing diversification and find:

Stylized fact 1: *Importers which are more geographically diversified in sourcing are less volatile.*

This can first be seen from Figure 1. Panel (a) plots a local polynomial regression of (log) firm level sales volatility on sourcing diversification measured by the average HHI between 2000 and 2006. Panel (b) looks at the volatility of exports and sourcing diversification. As can be seen, there is a general upward sloping trend in both figures: firms with more diversified sourcing

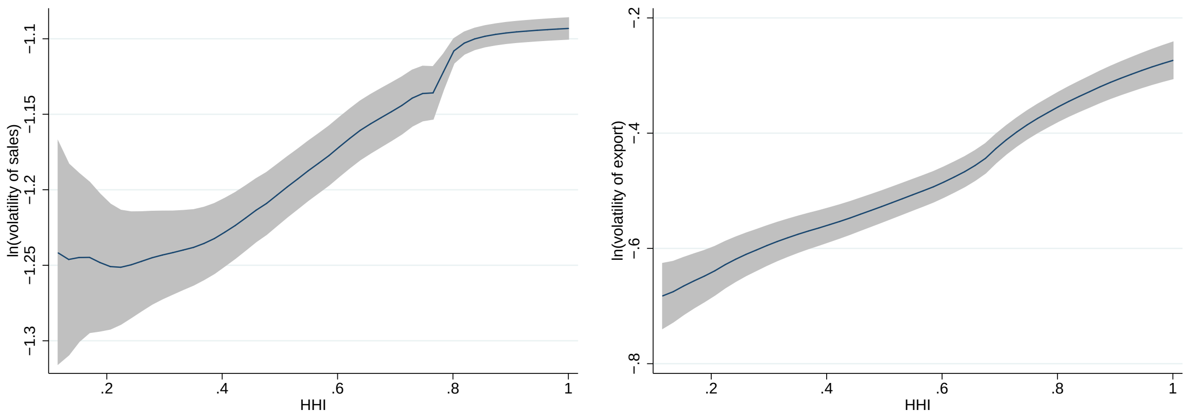
[14] In the Chinese customs regulations, importers are required to report the border customs district through which goods are actually imported. For the goods transferred between customs districts, the name of the customs district at the entry point is reported. For more details, please refer to section III of the Chinese *Standards on Completion of Customs Declaration Forms for Import/Export Goods*.

[15] The Chinese government further divides these provincial customs districts into more than 900 administrative areas. However, the customs data only report customs at the provincial level.

[16] Many of these custom districts were determined during the colonial era when the Chinese customs service was managed by the British. For example, the Shantou customs district used to be called *Chao Haiguan* and was established in 1860 to manage trade in the east of Guangdong province.

[17] The price index is from Brandt et al. (2012). I focus on the balanced panel and exclude entry and exit to insure that I have relatively longer time series to compute volatility.

[18] There is no product level price index for exports. Instead, I use output price index to deflate exports. The results are similar without deflating.



(a) volatility of sales

(b) volatility of exports

Figure 1: Sourcing diversification and firm-level volatility

strategies are associated with lower volatility. Of course, there may be confounders that lead to such a relationship. To handle such concern, I conduct a regression analysis regressing firms' output volatility on sourcing diversification, controlling for age, size in terms of average employment, and productivity measured in terms of average TFP during sample period for firms. I also control for diversification in the product margin by controlling the average number of imported products (Harmonized System 8 digit product), and geographically diversification on the demand side by adding the number of exporting routes used by the exporters. The results are shown in Table A1. The relationship remains stable: a higher HHI is associated with higher output volatility. It continues to hold when restricting the sample to importers and controlling for the number of products imported. The regressions on exports volatility lead to the same conclusion as in the Appendix Table A2.

2.2 Customs District Heterogeneity and Gravity

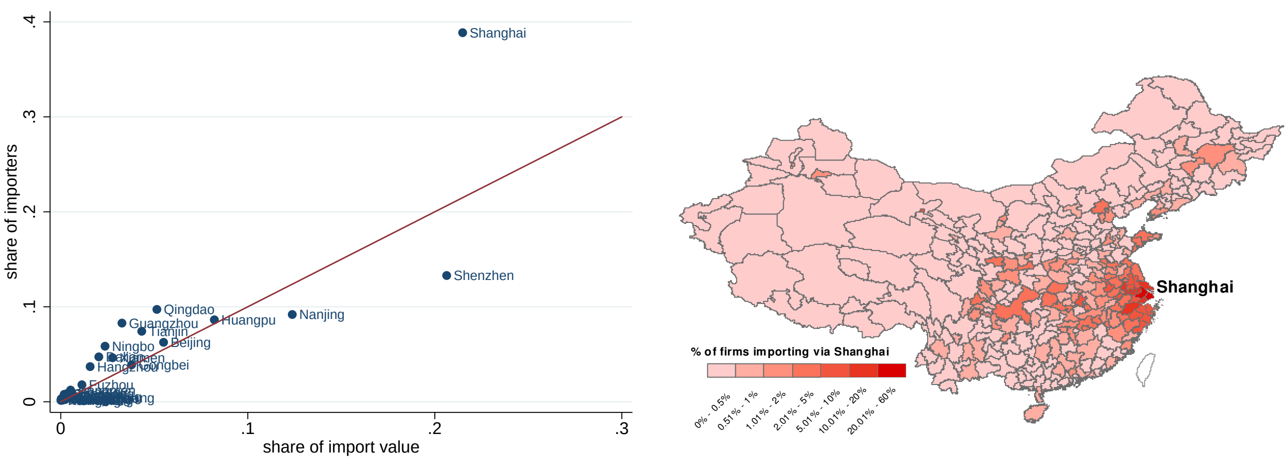
Importers source inputs through geographically distinct customs districts. These customs districts show rich heterogeneity in terms of the number of firms they serve and the value of goods they process. This is captured in Figure 2 (a). The figure plots the share of Chinese imports through each customs district on the horizontal axis against the share of importers that import via each customs district on the vertical axis.[19] The vertical axis captures the extensive margin while the horizontal axis captures both the intensive margin and the extensive margin. As is

[19]The sum of the values on the vertical axis does not add up to 1 because firms could import through multiple customs districts. The current result uses data from the year 2006 - results from other years are similar.

obvious from the figure, there are large variations between customs districts. The Shanghai customs district is the largest. Nearly 40% of Chinese importers import through Shanghai. Although the share of importers through Shenzhen is just about a third of Shanghai, the value of goods passing through is almost the same, about 20% of the total. Such divergence in the extensive margin and the intensive margin suggests that some customs districts may be easy to access but they are not as efficient in terms of sourcing foreign imports.[20]

Part of Shanghai's advantage is probably its relatively central position on the Chinese coastline. Figure 2 (b) plots the share of firms importing via Shanghai for each prefecture city. Gravity clearly plays a role: there is a gradient originating from Shanghai. Closer firms are more likely to import through Shanghai. These findings can be summarized as:

Stylized fact 2: *Customs districts are heterogeneous in facilitating imports and firms tend to source via closer customs.*



(a) Customs district heterogeneity

(b) Customs district gravity

Figure 2: Customs district heterogeneity and gravity

2.3 Multi-customs-district Premium

Firms using different numbers of customs districts are also very different. Figure 3 (a) shows the distribution of customs use across importers. Importers using multiple customs districts are a minority but they import much more goods than single-customs-district importers. Only 30%

[20]Of course, firms importing through each customs district, the trade costs with their trading partners, and their trading partners themselves might also be very different. Below, I will estimate the efficiency of each customs district taking these factors into account.

of the importers import via more than one customs district. But they contribute about 60% of total imports. This suggests that the importers using multiple customs districts are probably larger. I next examine whether this is borne out by regression analysis.

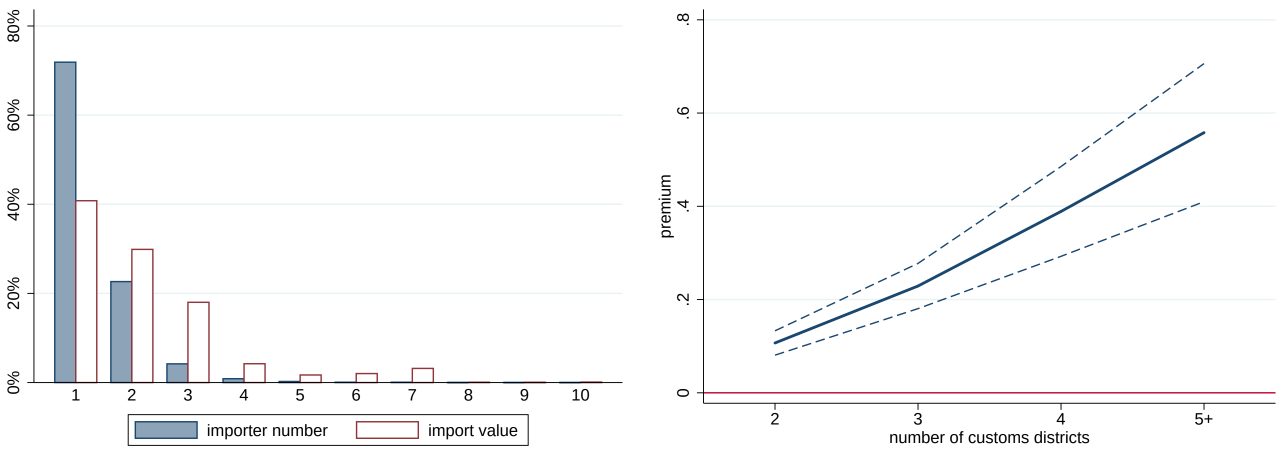
It is well known that importers are larger than non-importers (Bernard et al. 2007, Kugler and Verhoogen 2009). AFT shows that the importer premium tends to rise with the number of countries that firms import from. I confirm this finding in the Chinese data and show that there is an additional premium: importers importing through more customs districts tend to be larger and more productive. This is shown in Table A3 in which I use data from the year 2006 and regress firm characteristics on the number of customs districts that firms use, controlling for the number of origins. My focus is the dummies indicating the number of customs districts that importers use with single-customs-district firms as the benchmark group. Columns (1) to (4) focus on sales. Column (1) controls only for industry, prefecture and ownership fixed effects, and the premium of multi-customs-district firms is huge. Moreover, it increases with the number of customs districts. When the number of importing countries is included in column (2), which is the focus of AFT, the effect shrinks by around two thirds. Firm size as measured by employment is included in column (3). To address the concern that the premium could be due to multi-plant firms located in multiple customs districts, a measure controlling for multi-plant firms is added in column (4). Either CAIS or the customs data do not report the number of plants. Since the customs data report the destination and origin for each transaction, I count the number of distinct domestic destination/origin locations for each Chinese exporter/importer. If firms have separate plants in each location, this place count measure can be used to control for multi-plant firms. Adding this multi-plant measure, the premium decreases slightly but remains sizeable and significant. Similar results hold for imports in column (5), labor productivity as measured by real value added per worker in column (6) and Total Factor Productivity (TFP) in column (7).[21] The premium for sales is visualized in Figure 3 (b), with the dash lines indicating the 95% confidence intervals. The third stylized fact is summarized as:

Stylized fact 3: *Multi-customs-district sourcing firms are larger and more productive.*

In Appendix 9.1.2, I conduct various robustness checks on the premium, including an alternative measure for a multi-plant firm, excluding processing importers who are subject to place-based policy such as processing trade zone, and excluding importers from Guangdong

[21] I use the price indexes from Brandt et al. (2012) to construct real value added and real capital stock. TFP is estimated using the Levinsohn and Petrin (2003) method.

province which is divided into seven customs districts. For all these robustness checks, the premium remains sizeable and highly significant. The premium is not particular to the year 2006 and found in data from other years as well.



(a) Import distribution

(b) Size premium in sales

Figure 3: Multi-customs-districts premium

3 Theoretical Framework

This section presents a model of global sourcing which reconciles the three stylized facts established in the previous section. More importantly, it provides theoretical predictions on sourcing diversification and resilience to adverse shocks on the supply chain, which will guide my empirical analysis. I introduce multiple domestic regions, domestic trade costs, and customs services into the model by Antràs, Fort and Tintelnot (2017). While countries are singletons and goods arrive at factory doors directly in their model, the new features are necessary to identify domestic regions and customs districts hit by SARS. They also allow me to investigate the role of domestic infrastructure.

3.1 Demand

There are I regions at home. In each region, the representative consumer's preference for the manufacturing final goods is given by the following CES utility function

$$U_i = (\int_{\varpi \in \Omega_i} q(\varpi)^{\frac{\sigma-1}{\sigma}} d\varpi)^{\frac{\sigma}{\sigma-1}},$$

where $\sigma > 1$ is the demand elasticity, and Ω_i is the set of final-good varieties available at region i. The demand for final goods at region i is determined by

$$q_i(\varpi) = B_i p_i(\varpi)^{-\sigma},$$

where $B_i \equiv \frac{1}{\sigma}(\frac{\sigma}{\sigma-1})^{1-\sigma} P_i^{\sigma-1} E_i$ is a region specific demand shifter; E_i and P_i are the local expenditure and price index, respectively; $p_i(\varpi)$ is the price of variety ϖ.

3.2 Production and Trade

The final-good producers compete in a monopolistically competitive market with free entry. They are endowed with a core productivity φ which is heterogeneous and drawn from a distribution $G_i(\varphi)$, $\varphi \in [\overline{\varphi}_i, \infty]$. Following Melitz (2003), such productivity is learned only after paying the fixed entry costs of f_{ei}. To produce the non-tradable final goods, firms assemble intermediate inputs which are sourced from intermediate input producers located in different origins. The bundle of intermediate inputs has a continuum of measure one and is assumed to have a symmetric elasticity of substitution ρ.[22]

While AFT assumes that the final-good producers trade directly with the intermediate input producers and there are no domestic trade costs, I assume that it requires customs services when sourcing the foreign inputs, and trade is also costly at home. The reason for making these assumptions is twofold. First, importers use services provided by the customs bureau at various stages of the transaction. Even services which are not directly provided by the customs bureau, such as searching for the right suppliers, translating documents, or making payments, are usually provided by intermediaries located in the vicinity of the customs bureau. The cost and efficiency of the service vary across customs districts, which can help explain the large customs district heterogeneity observed in the second stylized fact.[23] Second, domestic trade costs are particularly high in developing countries. Atkin and Donaldson (2015) estimate that the distance elasticity for domestic trade costs is four to five times larger in Ethiopia or Nigeria than in the US. In the case of China, as pointed out by Young (2000), interregional competition leads to severe market segregation. It is important to understand how domestic trade costs shape firms' sourcing behaviour and how improvement in infrastructure might help firms in sourcing.

[22] The assumption of non-tradable final goods is not crucial. I relax this assumption later in one of the robustness checks. The measure of intermediate inputs can also be endogenized without changing the main predictions. Similar to AFT, ρ turns out to play little role in the model.

[23] Customs broker is a typical type of intermediary. Alibaba, the largest Chinese B2B platform, provides an online platform for customs brokers. The prices and services listed vary across locations.

To keep the model as tractable as possible and at the same time retaining these additional features, I assume that firms' input sourcing follows a two-stage process, as illustrated in Figure 4. Intermediate inputs are first sourced by intermediaries located in each customs district. Inputs are then shipped to the final-good producers.[24] The iceberg trade costs of shipping inputs from origin k to the customs district j, and from j to the final destination i are denoted as τ_{jk} and τ_{ij}, respectively. In order to source inputs through trade route jk, the final-good producers from region i need to pay a fixed cost in terms of f_{ijk} units of labor in region i. I use $J_i(\varphi)$ to denote the set of customs districts, and $K_{ij}(\varphi)$ the set of origins for which the firm with productivity φ located in region i has paid the associated fixed cost of sourcing $w_i f_{ijk}$. I will refer $J_i(\varphi)$ and $K_{ij}(\varphi)$ as the *sourcing strategy*.

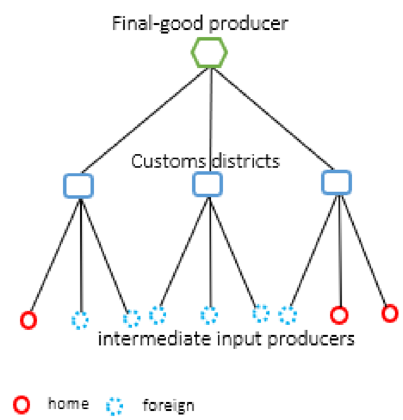


Figure 4: Illustration of firms' sourcing process

The intermediate input producers use constant return to scale technologies for production with labor as the only input, and sell their outputs competitively. At each origin, there is a continuum of intermediate input producers. The unit labour requirement is denoted as $a_k(\varphi, v)$ for the input producer $v \in [0, 1]$ locating in region k who supplies inputs for a firm with productivity φ. Following AFT, I assume that the firm-specific $a_k(\varphi, v)$ is drawn from the following Fréchet distribution:

$$\Pr(a_k(\varphi, v) > a) = e^{-T_k a^{\theta}},\ T_k > 0,$$

where T_k is the average efficiencies of intermediate input producers from origin k. At each customs district, there is a continuum of intermediaries which use constant return to scale technologies providing the customs service. The unit labor requirement for the intermediary

[24]This is similar to the "hub and spoke" structure used by Head, Jing and Ries (2017) for sourcing. It also resembles the idea of international gateway in Cosar and Demir (2016).

$\omega \in [0,1]$ locating in customs district j trading with the firm having productivity φ is denoted as $b_j(\varphi,\omega)$. Again, it is assumed that $b_j(\varphi,\omega)$ is drawn from a Fréchet distribution:

$$\Pr(b_j(\varphi,\omega) > b) = e^{-A_j b^\theta}, A_j > 0,$$

where A_j is the average efficiency of the intermediaries in customs district j. Under these assumptions, the marginal cost of firms which is given by

$$c_i(\varphi) = \frac{1}{\varphi}(\int_0^1 [\tau_{ij} b_j(\varphi,\omega) w_j (\int_0^1 (\tau_{jk} a_k(\varphi,v) w_k)^{1-\rho} dv)^{\frac{1}{1-\rho}}]^{1-\rho} d\omega)^{\frac{1}{1-\rho}}.$$

3.3 Optimal Sourcing

The final-good producers' problem in sourcing has two layers: the sourcing strategy, i.e., the extensive margin problem in choosing which trade routes to be used in sourcing inputs, and the intensive margin, i.e., how much inputs to source from each route. I first solve the intensive margin problem for a given sourcing strategy, then characterize the optimal sourcing strategy.

The cost of sourcing input v from k through intermediary ω at customs district j to destination i for firm φ is: $\tau_{ij}\tau_{jk}a_k(\varphi,v)b_j(\varphi,\omega)w_k w_j$. If I assume that the final-good producer learns about $a_k(\varphi,v)$ and $b_j(\varphi,\omega)$ simultaneously and seeks to $\min_{j,k}\{\tau_{ij}\tau_{jk}a_k(\varphi,v)b_j(\varphi,\omega)w_k w_j\}$, there is no explicit solution as in the Eaton-Kortum (2002) model. This is because the multiplication of two Fréchet distributed random variables is not Fréchet distributed. To make progress, I impose the the following assumption on timing: the final-good producers do not observe the realized unit labor requirement at the origins when making the sourcing decision across customs districts; they can only predict these costs given the productivity distribution of potential suppliers in different origin countries.[25] Suppose the expected unit cost of intermediate inputs shipped to customs district j for firm φ is $c_i^j(\varphi)$. The customs district picked by final-good producer is determined by solving the following problem:

$$\min_{j \in J_i(\varphi)} \{\tau_{ij} b_j(\varphi,\omega) c_i^j(\varphi) w_j\}.$$

Since $1/b_j(\varphi,\omega)$ is Fréchet distributed, according to Eaton and Kortum (2002), the probability

[25] Antràs and de Gortari (2017) make a similar assumption in a model of global value chain with multi-stage production. They show that this assumption of incomplete information with stage specific randomness is isomorphic to an alternative assumption of complete information but with randomness ascribed to the *overall costs* of a given route.

of sourcing through customs district j is given by

$$\chi_{ij}(\varphi) = \frac{A_j(\tau_{ij} w_j c_i^j(\varphi))^{-\theta}}{\sum_{l\in J_i(\varphi)} A_l(\tau_{il} w_l c_i^l(\varphi))^{-\theta}}. \tag{3.1}$$

The problem at customs district j in choosing intermediate input producers across origins is:

$$\min_{k\in K_{ij}(\varphi)} \{\tau_{jk} a_k(\varphi, v) w_k\}.$$

Again, given the Fréchet distributed $1/a_k(\varphi, v)$, the probability of sourcing from region k at customs district j is given by

$$\chi_{k|j}(\varphi) = \frac{T_k(\tau_{jk} w_k)^{-\theta}}{\Theta_j(\varphi)},$$

where $\Theta_j(\varphi) \equiv \sum_{n\in K_{ij}(\varphi)} T_n(\tau_{jn} w_n)^{-\theta}$. The expected unit cost $c_i^j(\varphi)$ is given by $c_i^j(\varphi) = (\gamma\Theta_j(\varphi))^{-\frac{1}{\theta}}$, where γ is a constant defined by the Gamma function. Similar to the Nested Logit model in discrete choice theory, the probability of sourcing from origin k using customs district j for final-good producer from region i with productivity φ, which I will call *sourcing intensity* later, is given by:

$$\chi_{ijk}(\varphi) = \chi_{ij}(\varphi)\chi_{k|j}(\varphi) = \frac{A_j T_k(\tau_{ij}\tau_{jk} w_j w_k)^{-\theta}}{\Psi_i(\varphi)}, \tag{3.2}$$

where $\Psi_i(\varphi) \equiv \sum_{l\in J_i(\varphi)} A_l \Theta_l(\varphi)(\tau_{il} w_l)^{-\theta} = \sum_{l\in J_i(\varphi), n\in K_{ij}(\varphi)} \phi_{iln}$ is the *sourcing capability* of the firm, and $\phi_{iln} = A_l T_n(\tau_{il}\tau_{ln} w_l w_n)^{-\theta}$ is the *sourcing potential* of origin n through customs district l. Then Equation (3.1) can also be rewritten as

$$\chi_{ij}(\varphi) = \frac{A_j \Theta_j(\varphi) \tau_{ij}^{-\theta} {w_j}^{-\theta}}{\Psi_i(\varphi)}.$$

Thus the customs districts which have lower costs trading with the destination are more likely to be used. This is consistent with Stylized fact 2 on customs district gravity. Following Eaton and Kortum (2002), the Fréchet assumptions implies that

$$c_i(\varphi) = \frac{1}{\varphi}(\gamma^2 \Psi_i(\varphi))^{-1/\theta}. \tag{3.3}$$

Up till now, the sourcing strategies given by $J_i(\varphi)$ and $K_{ij}(\varphi)$ have been taken as given.

They are characterized by the following problem:

$$\max_{I_{ijk}\in\{0,1\}_{j=1,k=1}^{J,K}} \pi_i(\varphi,\{I_{ijk}\}_{j=1,k=1}^{J,K}) = B_i\varphi^{\sigma-1}(\gamma^2\sum_{j=1,k=1}^{J,K} I_{ijk}A_jT_k(\tau_{ij}\tau_{jk}w_jw_k)^{-\theta})^{\frac{\sigma-1}{\theta}} - w_i\sum_{j=1,k=1}^{J,K} I_{ijk}f_{ijk}, \tag{3.4}$$

where I_{ijk} is an indicating variable, J and K are the total number of customs district and origins that firms could potentially choose. I_{ijk} takes value 1 if $j \in J_i(\varphi)$ and $k \in K_{ij}(\varphi)$, that is $J_i(\varphi) \equiv \{j : I_{ijk} = 1\}$ and $K_{ij}(\varphi) \equiv \{k : I_{ink} = 1, n = j\}$. As noted by AFT, there is no explicit solution to Problem (3.4). A brute force approach requires an evaluation of 2^{JK} combinations of customs district and origin for each firm. Nonetheless, the solution has the following properties.

Proposition 1. *The optimal sourcing strategy $I_{ijk}(\varphi) \in \{0,1\}_{j=1,k=1}^{J,K}$ is such that*

(a) a firm's sourcing capability $\Psi_i(\varphi)$ is non-decreasing in φ;

(b) if $\sigma - 1 > \theta$, $J_i(\varphi_L) \subseteq J_i(\varphi_H)$, $K_{ij}(\varphi_L) \subseteq K_{ij}(\varphi_H)$ for $\varphi_H \geq \varphi_L$;

(c) if $\sigma - 1 > \theta$, $\Theta_j(\varphi)$ is non-decreasing in φ.

Proof. See Appendix 8.1. □

Conclusion (a) implies that firms with higher core productivities φ have even lower marginal costs given their higher sourcing capabilities. In the case that $\sigma - 1 > \theta$, inputs are complementary. According to conclusion (b), there is a pecking order in firms' sourcing strategies.[26] It implies that high productivity firms are more likely to source not only from more origins but also via more customs districts. This is consistent with Stylized fact 3 that multi-customs-district importers are more productive.

3.4 Industry and General equilibrium

Following AFT, I assume that consumers spend a fixed share of their income η on the manufacturing final goods. The remainder is spent on an outside good which is homogeneous and freely tradable across regions. The outside good thus serves as numeraire and pins down the wage for each region. Wages are thus taken as given in solving the sectoral equilibrium for the

[26]For the case that $\sigma - 1 = \theta$, the sourcing decisions across different sourcing routes are independent. For the case that $\sigma - 1 < \theta$, the inputs are substitutable. In both cases, the sourcing strategies of firms do not necessarily follow a pecking order. For a more detailed discussion, please refer to AFT. In the rest of the paper, I focus on the more empirically relevant case that inputs are complementary. Later, I provide an estimate for $\sigma - 1 - \theta$ which turns out to be positive.

manufacturing sector. Since entry is free:

$$\int_{\overline{\varphi}_i}^{\infty} \pi_i(\varphi) dG(\varphi) = w_i f_{ei},$$

the measure of final-good producers in each region can be pinned down as

$$N_i = \frac{\eta L_i}{\sigma(\int_{\overline{\varphi}_i}^{\infty} \sum_{j \in J_i(\varphi), k \in K_{ij}(\varphi)} f_{ijk} dG_i(\varphi) + f_{ei})}.$$

3.5 The Gravity Equation

For a firm with productivity φ, if it sources inputs from origin k via customs district j, the corresponding import is given by

$$M_{ijk}(\varphi) = (\sigma - 1) B_i \varphi^{\sigma-1} (\gamma^2 \Psi_i(\varphi))^{\frac{\sigma-1}{\theta}} \chi_{ijk}(\varphi). \tag{3.5}$$

Then the total imports of all firms in region i from origin k via customs district j is given by

$$\begin{aligned} M_{ijk}(\varphi) &= N_i \int_{\overline{\varphi}_{ijk}} M_{ijk}(\varphi) dG(\varphi) \\ &= (\sigma - 1) B_i \gamma^{\frac{2(\sigma-1)}{\theta}} A_j T_k (\tau_{ij} \tau_{jk} w_j w_k)^{-\theta} \Lambda_{ijk}, \end{aligned}$$

where $\Lambda_{ijk} = \int_{\overline{\varphi}_{ijk}} I_{ijk}(\varphi) \varphi^{\sigma-1} \Psi_i(\varphi)^{\frac{\sigma-1-\theta}{\theta}} dG(\varphi)$ and $\overline{\varphi}_{ijk}$ is the productivity cut-off for firms located in region i picking route jk.

3.6 Diversification, Resilience, and Volatility

The previous results are direct extensions of AFT, this subsection presents *new* results on firms' diversification in sourcing, resilience to shocks on supply chains, and output volatility. Proposition 1 implies that high productivity firms tend to be more diversified along the extensive margin since they source from more trade routes. However, it is not necessarily true that their inputs are less concentrated. For example, suppose firm A is using two trade routes with each contributing $\frac{1}{2}$ of total inputs, while firm B is using three routes with one contributing $\frac{3}{4}$, and the other two each contributing $\frac{1}{8}$. The concentration of A's sourcing strategy measured by the HHI is $(\frac{1}{2})^2 + (\frac{1}{2})^2 = \frac{1}{2}$, and $(\frac{3}{4})^2 + 2(\frac{1}{8})^2 = \frac{19}{32} > \frac{1}{2}$ for B. So B looks more diversified by the extensive margin, but less diversified after taking the intensive margin into account. The following proposition rules out such a possibility.

Proposition 2. *Under input complementarity $\sigma - 1 > \theta$, the concentration of firms' sourcing strategies as measured by the Herfindahl-Hirschman Index $HHI_i(\varphi) \equiv \sum_{j,k} \chi_{ijk}(\varphi)^2$ is non-increasing in φ.*

Proof. See Appendix 8.2. □

Therefore, high productivity firms are more diversified even after considering the intensive margin. The intuition is that if a certain sourcing route is dominant for a firm, it must be less or equally dominant for a more productive firm. This is because the high productivity firms have greater sourcing capability and more alternatives. For the example above, it cannot be that B's most dominant option takes a share greater than $\frac{1}{2}$ when it has one more option than A.

So far, I have characterized the properties of the optimal sourcing strategy for forgiven sourcing potentials and fixed costs. The following proposition considers a comparative statics on how the optimal source strategies respond to exogenous changes in these parameters.

Proposition 3. *Under input complementarity $\sigma - 1 > \theta$ and fixed market demand B_i, firms' sourcing strategies $J_i(\varphi)$ and $K_{ij}(\varphi)$ weakly expand whenever there is improvement in the sourcing potential $\vec{\phi}_i$ or reduction in the fixed costs of sourcing $\vec{f}_i$, where $\vec{\phi}_i{}' = \{\phi_{ijk}\}_{j=1,k=1}^{J,K}$ and $\vec{f}_i{}' = \{f_{ijk}\}_{j=1,k=1}^{J,K}$.*

Proof. See Appendix 8.3. □

The proposition implies that increasing sourcing potentials or reduction in the fixed costs of sourcing will induce firms to expand their sourcing strategies along the extensive margin.[27] The intuition behind the result is as follows. Since inputs are complementary, an increase of sourcing potential of any sourcing route is likely to raise the marginal benefit of including a route in the sourcing strategy. Reducing the cost of any sourcing route is likely to lower the marginal cost of including a route. These make it more likely for a firm to add a new route.

Now I examine the model's prediction on firms' resilience to shocks. Resilience is measured by the pass-through of adverse shocks to firm performance. A firm is said to be more resilient if the pass-through is smaller. Since there is no explicit solution to the model, we might expect that it is difficult to know without numerical simulations. It turns out that we can gauge the effect without solving the whole model by using the "hat algebra" technique thanks to Jones (1965) and revitalized by Deckle, Eaton, and Kortum (2007). One complication is that my

[27] In a different model setup, Bernard et al. (2015) discovers a similar result with respect to search costs and variable trade costs.

model has adjustments in the extensive margin. Firms can add or drop sourcing routes. In the Eaton-Kortum typed models, firms import from everywhere and there is only adjustment in the intensive margin. To solve the problem, I use a technique from Feenstra (1994) with which he estimates the welfare gains from new varieties. Applying his idea along with the hat algebra approach, I find:

Proposition 4. *For a small idiosyncratic trade cost shock which changes τ_{ijk} to τ'_{ijk} ($\tau_{ijk} \equiv \tau_{ij}\tau_{jk}$) such that the firm does not abandon route jk, we have:*

(a) The pass-through to the marginal cost is given by

$$\frac{\partial \ln(\widehat{c_i(\varphi)})}{\partial \ln(\widehat{\tau_{ijk}})} = \frac{\chi_{ijk}(\varphi)}{1 - \sum_{jk \in \mathcal{N}_i(\varphi)} \chi'_{ijk}(\varphi)},$$

where $\widehat{X} \equiv \frac{X'}{X}$ and $\mathcal{N}_i(\varphi)$ is the set of new routes chosen by the firm after the shock.

(b) Under input complementarity ($\sigma - 1 > \theta$) and adverse shock ($\tau'_{ijk} \geq \tau_{ijk}$),

$$\frac{\partial \ln(\widehat{c_i(\varphi)})}{\partial \ln(\widehat{\tau_{ijk}})} = \chi_{ijk}(\varphi),$$

$$\frac{\partial^2 \ln(\widehat{c_i(\varphi)})}{\partial \ln(\widehat{\tau_{ijk}})\partial\varphi} \leq 0; \frac{\partial^2 \ln(\widehat{c_i(\varphi)})}{\partial \ln(\widehat{\tau_{ijk}})\partial\phi_{ijk}} > 0.$$

That is, high productivity firms are more resilient to adverse shocks, and firms are less resilient to shocks on more appealing sourcing routes.

Proof. See Appendix 8.4. □

The pass-through has two components according to conclusion (a): the intensive margin captured by $\chi_{ijk}(\varphi)$ and the extensive margin captured by $\frac{1}{1-\sum_{jk\in\mathcal{N}_i(\varphi)} \chi'_{ijk}(\varphi)}$. Both depend on firm productivity φ.[28] However, it is difficult to know how pass-throughs vary with productivity φ for general shocks. Conclusion (b) instead focuses on adverse shocks which is more relevant to the discussion of resilience.[29] In this case, the pass-through depends only on the intensive margin. This is because no firms will add *new* sourcing routes facing adverse shocks, according to Proposition 3. The only possible adjustment along the extensive margin is to drop sourcing

[28] The pass-through depends only on the intensive margin and is homogeneous across all importers in Eaton-Kortum typed models with universal importing. Therefore, these models predict that firms are equally resilient.

[29] The bottleneck problem is that we cannot characterize how the set of new routes $\mathcal{N}(\varphi)$ varies with φ. For a favorable shock, the pass-through is non-monotonic with respect to firm productivity which I discuss in the proof.

routes.[30] Then the term on extensive margin adjustment becomes $\frac{1}{1-\sum_{jk\in\mathcal{N}_i(\varphi)}\chi'_{ijk}(\varphi)} = 1$ since $\sum_{jk\in\mathcal{N}_i(\varphi)}\chi'_{ijk}(\varphi) = 0$. The impact of the shock is determined by the intensive margin and increases with $\chi_{ijk}(\varphi)$. If the firm is not diversified at all, and solely replies on the sourcing route hit by the shock, the pass-through is 100%. Conclusion (b) tells us that the pass-through decreases weakly with firm productivity. This is because high productivity firms are more diversified and source from more places. Their load of inputs on any particular route is smaller, and so is the pass-through. It also tells us that the pass-through is larger for routes with higher sourcing potential. Due to the pecking order, firms agree on the ranking of sourcing routes. The more appealing routes take larger shares for every firm. Shocks on these routes have higher pass-throughs and are more detrimental to firms.

Marginal costs are usually not directly observable. To generate empirically testable predictions, I study how the easily observed firm-level import flows given by Equation (3.5), will respond to an adverse shock. The model delivers the following result.

Proposition 5. *(a) For a small trade cost shock which increase τ_{imn} to τ'_{imn} such that firms do not abandon route mn, import flows respond according to*

$$-\frac{\partial \ln \widehat{M}_{ijk}(\varphi)}{\partial \ln \widehat{\tau}_{imn}} = \begin{cases} \theta + (\sigma - 1 - \theta)\chi_{imn}(\varphi), & \text{if } m{=}j,\ n{=}k, \\ (\sigma - 1 - \theta)\chi_{imn}(\varphi), & \text{otherwise.} \end{cases}$$

(b) Under input complementarity, the size of the pass-through to imports decreases weakly with firm productivity.

Proof. See Appendix 8.5. □

Again, the pass-through endogenously depends on firm productivity φ. Other than the usual Fréchet shape parameter θ which captures the direct impact of the shock, there is an additional term $(\sigma-1-\theta)\chi_{imn}(\varphi)$ which is positive if inputs are complementary ($\sigma - 1 > \theta$), and negative if inputs are substitutable ($\sigma - 1 < \theta$). This additional term highlights the interdependencies across sourcing routes and disappears in the knife-edge case of no interdependencies ($\sigma - 1 = \theta$). The cost shock reduces firms' sourcing capability and increases their marginal cost according to Proposition 4. This drives down marginal demand curve for all inputs if inputs are complementary. Such a feedback effect through interdependencies amplifies the initial cost shock and reduce imports further. In contrast, if the inputs are substitutable, the cost shock reduces firm

[30] The situation here is to the opposite of Proposition 3. When the sourcing potential of a certain route declines, firms' sourcing strategies either shrink or remain the same.

output and drives up the marginal demand curve. Such increase in the marginal demand for the input dampens the initial negative shock. This difference will allow me to identify if inputs are complementary or substitutable.

The pass-through also varies the sourcing intensity $\chi_{ijk}(\varphi)$. The feedback effect is stronger if the firm has a heavier load on inputs from the route being shocked. Since $\chi_{ijk}(\varphi)$ decreases weakly with φ, the feedback effect is smaller for high productivity firms. Finally, the interdependency is also reflected by the result that imports also respond to shocks on other routes in the firm's sourcing strategy.

I have shown that more productive firms can be more resilient to adverse shocks in Proposition 4. However, since high productivity firms are sourcing from more places, they may also be more exposed to shocks. There is no guarantee that they are less volatile. The following proposition provides conditions under which high productivity firms are also less volatile.

Proposition 6. *(a) If the shocks on sourcing routes are not perfectly correlated and have the same variance ξ^2, opening to trade lowers the volatility of firms' sourcing capabilities.*

(b) Under input complementarity, if the adverse shocks across sourcing routes are i.i.d., the volatility of firm revenue is:

$$var(\widehat{R_i}(\varphi)) \propto \xi^2 HHI_i(\varphi)$$

which weakly decreases with productivity φ.

(c) The volatility of importers is the same under universal importing.

Proof. See Appendix 8.6. □

While the literature has shown extensively that trade in intermediate brings productivity gains for firms (Goldberg et al., 2010; Halpern et al., 2015; Blaum et al., 2016; AFT, 2017), we know less about the effect on higher moments of firm performance and how they vary with firm productivity. Result (a) indicates a potential additional benefit of opening to trade for intermediate inputs: lower firm-level volatility. Caselli et al. (2015) illustrate that opening to trade can lower countries' aggregate volatility by allowing countries to diversify and reducing the exposure to domestic shocks. A similar mechanism is present in my model at the firm level except that I allow firms to add or drop sourcing routes while countries import from everywhere in their model. They emphasize that the mechanism hinges on the size of variance and covariance across countries. This is still true in my model. If the variance of domestic sourcing potential is negligible compared with the variance of the foreign sourcing potentials, sourcing autarky actually would lead to lower volatilities.[31] Result (b) spells out a scenario that more geographically diversified firms are less volatile. It provides a theoretical explanation for Stylized fact 1. The channel relies on diversification since the variation of volatility is all loaded on the variation of HHI. In a model with universal importing such as Eaton-Kortum, result (c) implies homogeneity in the volatility of all importers, regardless of the underlying structure of shocks. This is at odds with Stylized fact 1 and highlights the importance of adjustment in the extensive margin in generating volatility heterogeneity across firms.

4 Diversification and Resilience to the SARS Epidemic

This section tests Proposition 5 on diversification and resilience by exploiting the 2003 SARS epidemic as a natural experiment. During my sample period, SARS was one of the most significant events which disrupted the supply chains of China.[32] As I will argue, it was an unexpected exogenous shock to Chinese firms which made it more difficult for them to source inputs.

4.1 The SARS Epidemic

SARS was the first easily transmissible epidemic to emerge in the new millennium. It broke out in Southern China in November 2002 and ended in July 2003. It was an unknown disease which has respiratory symptoms similar to an influenza and could not be cured by existing antivirals and antibiotics. Given its severity and infectiousness, governments and intergovernmental organizations took unprecedented measures to prevent it becoming a global pandemic (Heymann et al. 2013). Other than travel advice warning people against travelling to areas with local outbreaks, WHO also issued procedures to hold cargo vessels in check at ports in case there were probable cases on board.[33] The International Civil Aviation Organization (ICAO) set up the "Anti-SARS Airport Evaluation Project" to impose checks on flights from SARS infected areas.[34] These measures necessarily created frictions in the flow of people and goods. For example, the number of air passengers around the Asia-Pacific dropped almost by 50% in the second quarter of 2003 compared with 2002 (Hollingsworth et al., 2006), while the freight traffic in Asia

[31] This may explain why Kurz and Senses (2016) find that US importers are more volatile than non-importers. As a matured economy, US is probably less volatile than other countries.

[32] While Japan is one of the key suppliers for China and there are many recent studies on the effect of the 2011 earthquake (Boehm et al 2015; Todo, et al., 2015; Carvalho et al., 2016), my data do not cover this period.

[33] For example, *Travel advice - Hong Kong Special Administrative Region of China, and Guangdong Province, China* was issued on 2 April 2003. *Procedures for prevention and management of probable cases of SARS on international cargo vessels* was issued on 23 May 2003.

[34] ICAO Airport Evaluation for Anti-SARS Protective Measures.

and North American stayed below the 2002 level for most of the year.[35]

Important suppliers of mainland China such as Canada, Hong Kong, Taiwan, Singapore and Vietnam were severely affected. These five regions topped the list of SARS cases, just behind mainland China itself (see Table A12). Imports from these regions alone made up about 20.5% of total Chinese imports in 2002. SARS struck these regions on different dates and lasted for different periods, such spatial and time variations help to identify the effect of the epidemic. To capture these variations, I use lists of *Areas with recent local transmission of SARS* provided by WHO which it identified as risky to travel to. These lists are summarized in Table A6, in which I indicate the period that each region is listed as risky. Using this list, I construct a dummy $SARS_{jk,t}$ indicating whether a trade route is hit by SARS or not at period t. It takes the value one as long as a Chinese customs district j or origin k remains on the list at time t.[36] Since the listing depended not only on the development of the epidemic but also the discretion of WHO, it is very likely to be exogenous to Chinese importers and their foreign suppliers.

4.2 The Resilience of Firms to SARS

Proposition 5 predicts that the effect of an adverse shock on imports varies with the pre-shock sourcing intensity. To capture such a differential treatment effect, I run the following regression:

$$\begin{aligned}\ln Import_{ijk}^{nt} &= D_i + C_j + O_k + F^{nt} + \sum_k b_k X_k^{nt} + \alpha_1 \chi_{ijk}^{n,t-1} \\ &+\alpha_2 SARS_{jk,t} + \beta \chi_{ijk}^{n,t-1} SARS_{jk,t} + \gamma CoSARS_{ijk}^{nt} + \epsilon_{ijk}^{nt}, \end{aligned} \tag{4.6}$$

in which I examine how firm n's imports flowing from origin k through customs district j at time t, $Import_{ijk}^{nt}$, would respond if trade route jk was hit by SARS. The customs data are at monthly frequency which I aggregate to quarter-level to deal with the lumpiness of monthly data. According to Proposition 5, the pass-through depends on the sourcing intensity before shock $\chi_{ijk}^{n,t-1}$, which is measured by the average expenditure share of firm n for inputs from route jk before the SARS epidemic. I add an interaction term between the SARS shock $SARS_{jk,t}$ and the pre-SARS sourcing intensity $\chi_{ijk}^{n,t-1}$ to capture the heterogeneous pass-through, while controlling the main effects. The main coefficient of interest is β. It has a structural interpretation of $-(\sigma - 1 - \theta)$ and is expected to be negative if inputs are complementary, but *vice versa* if inputs are substitutable. I also control for the interdependence of trade flows across routes by adding a

[35] IATA International Traffic Statistics: December 2003.

[36] A customs district is defined as affected if its local province is on the list.

dummy $CoSARS_{ijk}^{nt}$ indicating whether other sourcing routes of firm n were hit by SARS or not at period t. This is because Proposition 5 predicts that trade flows also respond to shocks hitting other routes. At the same time, I control for D_i the destination fixed effect, C_j the customs district fixed effect, and O_k the origin fixed effect, respectively. Finally, and most importantly, I control for firm characteristics X_k^{nt} and firm-time fixed effects F^{nt} to deal with idiosyncratic firm-level time-varying shocks, including demand shocks and disruptions in production due to the SARS epidemic.

The results are presented in Table 1, where I add the independent variables one by one. Unsurprisingly, the pre-shock sourcing intensity is positive and highly significant in all columns; trade flow is larger for routes with higher sourcing intensity. The main effect of the SARS shock is negative and highly significant as indicated in column (2). Firm imports fall by 7.9% on average if the route is hit by SARS. However, there are significant variations across firms and routes. The pass-through is much smaller for more diversified firms. When the interaction term between the sourcing intensity and SARS shock is introduced in column (3), the main effect of the SARS shock is reduced to about 5.6%, and the coefficient of the interaction term is negative and highly significant at -0.465. Then for a firm without any diversification before the epidemic such that $\chi_{ijk}^{n,t-1} = 1$, the overall effect of the SARS shock is $-0.0555 + (-0.465) * 1 = -0.52$. That is, its imports fell as much as 52%. In contrast, if the firm is very diversified such that $\chi_{ijk}^{n,t-1} \simeq 0$, the overall effect the SARS shock would just be the main effect at -5.6%. Moreover, the fact that the estimated β is negative implies $\sigma - 1 > \theta$, and the inputs are complementary. In column (4), I further include the dummy indicating whether other routes were hit by SARS to capture the interdependence across trade routes. The effect turns out to be very small and not significantly different from zero while the other coefficients remain robust.

4.3 Robustness Checks

4.3.1 Export Demand Shocks

So far, I have assumed that firms do not export, but in fact many importers are simultaneously exporters. If export demand shocks due to the SARS epidemic correlate with import cost shocks, the omitted variable problem will lead to a bias in the estimation. To understand how export demand shocks translate into import demand, I extend the model and allow firms to export final goods following the Melitz (2003) setup. This is shown in Appendix 8.8. The key result is that

Table 1: Resilience of firms to the SARS epidemic

Dependent Variable:	firm imports by route $ln(imp_{ijk,t})$			
	(1)	(2)	(3)	(4)
pre SARS sourcing intensity	9.461***	9.461***	9.493***	9.493***
	(0.0970)	(0.0970)	(0.0953)	(0.0953)
trade route hit by SARS=1		-0.0794***	-0.0555**	-0.0557**
		(0.0229)	(0.0244)	(0.0241)
trade route hit by SARS=1 x pre SARS sourcing intensity			-0.465***	-0.464***
			(0.133)	(0.133)
other routes hit by SARS=1				0.00454
				(0.0197)
firm-time FE	Y	Y	Y	Y
industry FE	Y	Y	Y	Y
ownership type FE	Y	Y	Y	Y
origin FE	Y	Y	Y	Y
destination FE	Y	Y	Y	Y
customs area FE	Y	Y	Y	Y
R^2	0.472	0.472	0.472	0.472
No. of observations	2019727	2019727	2019727	2019727

Notes: A trade route is a combination of an origin and a customs district. It is defined as hit by SARS if the origin or the customs district is listed by the WHO as regions with local transmission of SARS. The pre shock sourcing intensity is constructed as the route-specific input expenditure share averaged before the SARS epidemic. The numbers in the parentheses are standard errors clustered at region-industry level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

the pass-through of a shock affecting both exports and imports to import flow is given by

$$-\frac{\partial \ln \widehat{M}_{ijk}(\varphi)}{\partial \ln \widehat{\tau}_{imn}} = \begin{cases} \theta + (\sigma - 1 - \theta)\chi_{imn}(\varphi) + (\sigma - 1)\mu_{imn}(\varphi), \text{ if m=j, n=k,} \\ (\sigma - 1 - \theta)\chi_{imn}(\varphi) + (\sigma - 1)\mu_{imn}(\varphi), \text{ otherwise,} \end{cases}$$

where $\mu_{imn}(\varphi)$ is the intensity of final goods exported through trade route mn, which captures the diversification on the demand side. The pass-through is smaller for a more diversified exporter who has a smaller share of goods exported along route mn. Moreover, if $cov(\chi_{imn}(\varphi), \mu_{imn}(\varphi)) > 0$ and $\sigma - 1 > \theta$, so that imports and exports are positively correlated, the effect of diversification in sourcing is overestimated when diversification on the demand side is omitted.[37] To control for export demand shocks, I follow the theory to add an interaction term between the epidemic shock and the pre-SARS export intensity for each route. The export intensity is constructed as the average share of outputs exported through each route before the epidemic. The results are presented in Table 2. Column (1) is the benchmark which only includes

[37] $\sigma - 1 > \theta$ naturally implies $\sigma - 1 > 0$ because θ is greater than zero.

the sourcing diversification channel. Column (2) instead only includes the export diversification channel and omits the sourcing diversification channel. The coefficient for the export intensity is positive. So indeed firms tend to import more through the trade routes which have higher export intensity. Moreover, the interaction term between the SARS shock and export intensity has a significant negative coefficient. Thus, without looking at sourcing diversification, it looks as if imports are more resilient when the export intensity is lower. However, only diversification on the import side matters when I put the two channels together in column (3): the magnitude of the interaction term between the SARS shock and export intensity drops dramatically and is not significantly different from zero. At the same time, the baseline result of sourcing diversification remains robust and significant.

Alternatively, we expect that importers who do not export should not be exposed to export demand shocks. In Appendix Table A8, I split the sample into exporters and non-exporters. For importers who do not export, the differential treatment effect remains robust and significant.[38]

4.3.2 No Pre-Trend Assumption

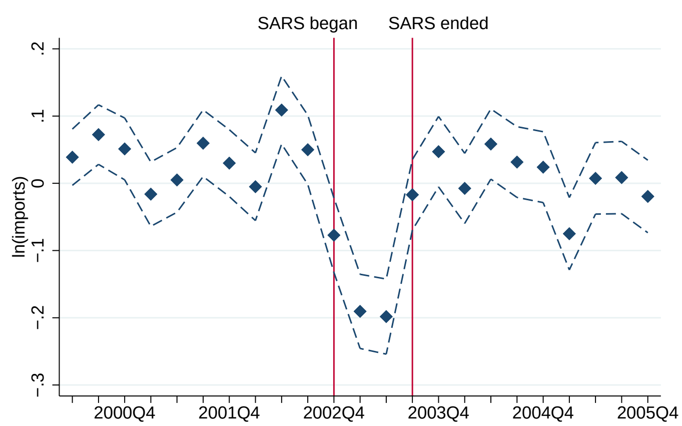


Figure 5: SARS on imports: difference in difference Estimation

Notes: The figure plots the coefficients of the interaction terms of the time dummy and the treated dummies in a difference-in-difference regression on firm imports by route. A firm-route is treated if either the importing origin or the customs district was affected by SARS during the sample period. Dash lines indicate 95% confidence intervals while standard errors are clustered at firm level.

Although I have controlled for a rich set of fixed effects and even firm-time fixed effects which

[38]The effect is much larger than the full sample. But these non-exporters are less productive than the exporters. They are less diversified and should have higher pass-throughs.

should alleviate much concern on selection. It might still be of concern that the routes hit by SARS were selected within a firm in such a way that made them more or less resilient to the SARS shock. To show that this is not the case and the SARS shock was as good as random, I employ a Difference-In-Difference strategy to estimate firm imports by route, and include the interaction terms of the time dummies and the treated dummy, controlling firm characteristics including firm size, age, and firm, industry fixed effects, ownership type, and origin-customs-destination fixed effects. A firm-route is defined as treated if it was eventually affected by SARS during the sample period. The coefficients for the interaction terms of the treated dummy and the time dummies are plotted in Figure 5. As we can see, there was no pre-existing trend before the epidemic.

4.3.3 Processing Importers

During the past three decades, China has adopted policies which encourage local firms to form processing trade relationships with foreign firms. Processing trade accounted for about half of total imports in the early 2000s (Yu, 2015). For Chinese processing traders, there are two important regimes (Feenstra and Hanson, 2005; Manova and Yu, 2016): processing with inputs (PI) under which firms independently source and pay for imported inputs, and pure assembly (PA) under which firms receive inputs at no cost from foreign partners.[39] As argued by Feenstra and Hanson (2005), and Manova and Yu (2016), PA firms play little part in sourcing and incur no costs using the imported inputs. Their sourcing decisions are at the discretion of their foreign partners. Moreover, the terms of the transaction with the foreign partner must be written in contracts and presented to the customs authority in advance. Given these institutional constraints, there is little scope for them to adjust their sourcing strategy in the face of unexpected shocks compared with normal firms. In contrast, PI firms take ownership of the imported inputs. They actively search for the right inputs and pay for the associated costs. Other than paying zero duties, the problem that PI firms face can still be described by my model. Thus their response to the SARS shock should still be in line with the model's prediction.

To see whether this is the case, I examine the effect of SARS on *pure* PI processing importers and *pure* PA processing importers separately. These are processing firms that only engage in processing imports.[40] The results are presented in Table 3. Columns (1) and (2) include only

[39] As long as the finished outputs are re-exported, both types of processing trade are exempted from import duties. If the processed goods are sold domestically, the exempted import tariffs must be returned.

[40] I also examine importers who partially participate in processing imports. The results are qualitatively the same as presented in Appendix Table A7. As noted by Yu (2015), there are hybrid firms which have both PI

Table 2: Resilience to SARS shock: including export demand shock

Dependent Variable: firm import by route $ln(imp_{ijk,t})$	import cost shocks only (1)	export demand shocks only (2)	both export demand and import cost shocks (3)
trade route hit by SARS=1	-0.0557**	-0.0696***	-0.0549**
	(0.0241)	(0.0244)	(0.0242)
pre SARS sourcing intensity	9.493***		9.470***
	(0.0953)		(0.0982)
trade route hit by SARS=1 x pre SARS sourcing intensity	-0.464***		-0.463***
	(0.133)		(0.132)
pre SARS export intensity		1.277***	0.0725**
		(0.0579)	(0.0362)
trade route hit by SARS=1 x pre SARS export intensity		-0.261***	0.0285
		(0.0741)	(0.0519)
other routes hit by SARS=1	0.00454	-0.0111	0.00467
	(0.0197)	(0.0224)	(0.0197)
firm-time, industry, ownership, origin, destination, customs FE	Y	Y	Y
R^2	0.472	0.396	0.472
No. of observations	2019727	2019284	2019284

Notes: A trade route is a combination of an origin and a customs district. It is defined as hit by SARS if the origin or the customs district are listed by the WHO as regions with local transmission of SARS. Pre shock export intensity is constructed as the average share of outputs exported through each route before the SARS epidemic. It is zero for non-exporters. The pre shock sourcing intensity is constructed as the input expenditure share averaged before the SARS epidemic. The numbers in the parentheses are standard errors clustered at region-industry level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

the sample of PI importers while columns (3) and (4) include only the sample PA importers. As expected, the response of PI firms is in line with the model. The coefficient of the interaction term is negative and significant, although the average effect is not significant. In contrast, the coefficient of the interaction term is positive but not significantly different from zero for PI firms. The main effect of the SARS is also not significantly different from zero. These results implies that the PA firms were not responsive in sourcing when affected by SARS, regardless their diversification in sourcing.

4.3.4 Alternative Cushioning Mechanisms

Finally, I examine whether the diversification channel under examination is robust to alternative mechanisms that might make firms more resilient to the SARS shock. The alternative channels that I consider include liquidity, finance and inventory. The idea is that firms with more liquidity, better access to credit or more inventory may also be more resilient to the SARS shock. These favorable conditions provide buffers for firms to absorb and counteract adverse shocks. If these firms at the same time are also more diversified, I would overestimate the effect of diversification. To rule out such a possibility, I construct measures to capture these various channels. Following Manova and Yu (2016), I measure the liquidity available to firms as (current assets - current liabilities)/total assets. For access to credit, I use the leverage ratio which is measured as liabilities/assets. Finally, for the inventory channel, I use the ratio of the inventory in intermediate inputs relative to total intermediate inputs. These variables are added as additional controls to the baseline regression. Since these measures are firm-year specific and will be fully absorbed by the firm-time fixed effect, the firm-time fixed effect is replaced by a county-time fixed effect. The results are presented in Table 4. Column (1) is the baseline which includes only the import diversification channel. Columns (2) to (4) focus on the alternative channels. Column (5) puts them together with the baseline channel. As we can see, these alternative channels do not appear have significant in cushioning the SARS shock on imports but the diversification channel remains significant.

There is concern that multi-plant firms, which have production in multiple locations and naturally import from more routes, are more resilient because of diversification in production. To deal with such concern, I focus on firms importing/exporting in a single location, which make up about 80% of the importers in my sample, and are likely to be single-plant firms. The results are presented in Appendix Table A9. The baseline result still holds for these firms

imports and PA imports. To make the test as clean as possible, these hybrid firms are excluded.

Table 3: Resilience of processing importers

Dependent Variable:	Processing with inputs			Pure Assembly		
firm import by route $ln(imp_{ijk})$	(1)	(2)	(3)	(4)	(5)	(6)
pre SARS sourcing intensity	7.744***	7.819***	7.819***	6.368***	6.364***	6.364***
	(0.189)	(0.192)	(0.192)	(0.383)	(0.378)	(0.377)
trade route hit by SARS=1	-0.0327	0.0238	0.0568	0.0438	0.0406	0.107
	(0.0499)	(0.0494)	(0.0655)	(0.148)	(0.158)	(0.162)
trade route hit by SARS=1 x pre SARS sourcing intensity		-0.648***	-0.643***		0.0423	0.0432
		(0.195)	(0.195)		(0.744)	(0.739)
other routes hit by SARS=1			0.0487			0.147
			(0.0616)			(0.144)
firm-time, industry, ownership, origin, destination, customs FE	Y	Y	Y	Y	Y	Y
R^2	0.530	0.530	0.530	0.529	0.529	0.529
No. of observations	267005	267005	267005	17556	17556	17556

Notes: Pure processing importers are firms which have no ordinary imports subject to tariffs. Columns (1) and (3) include the sample of importers which only engage in processing with inputs (PI). Column (4) and (6) include the sample of importers only engaged in pure assembly(PA). PA firms do not decide where to source or pay for the sourced inputs while PI firms do. The numbers in the parentheses are standard error clustered at region-industry level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table 4: Robustness check: the liquidity, finance, and inventory channels

Dependant Variable:	firm imports by route $ln(imp_{ijk,t})$				
	(1)	(2)	(3)	(4)	(5)
pre SARS sourcing intensity	9.124***	9.124***	9.124***	9.138***	9.141***
	(0.0900)	(0.0900)	(0.0900)	(0.0902)	(0.0906)
trade route hit by SARS	-0.0404**	-0.0391*	-0.0391*	-0.0405**	-0.0105
	(0.0197)	(0.0201)	(0.0227)	(0.0206)	(0.0375)
trade route hit by SARS x pre shock sourcing intensity	-0.321***	-0.320***	-0.321***	-0.321***	-0.312***
	(0.0955)	(0.0954)	(0.0953)	(0.0956)	(0.0954)
liquidity		0.00681			-0.0840**
		(0.0201)			(0.0366)
trade route hit x liquidity		-0.0101			-0.0568
		(0.0277)			(0.0586)
leaverage ratio			-0.0131		-0.0435***
			(0.00860)		(0.0153)
trade route hit x leverage ratio			-0.00151		-0.0237
			(0.0138)		(0.0289)
inventory				-0.324***	-0.329***
				(0.0179)	(0.0192)
trade route hit x inventory				0.000131	-0.0162
				(0.0286)	(0.0310)
firm, destination-time, ownership, industry, origin, customs FE	Y	Y	Y	Y	Y
R^2	0.403	0.403	0.403	0.403	0.403
No. of observations	2143515	2143515	2143515	2143515	2140168

Notes: Following Manova and Yu (2016), liquidity available to firms is measured by (current assets - current liabilities)/total assets. Inventory is measured as the ratio of intermediate inputs relative to total inputs. In all regressions, we control time, destination region, origin, customs region and ownership fixed effects. The numbers in the parentheses are standard errors clustered at region-industry level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

importing/exporting in a single place, while the effect is not significant for firms with multiple importing/export locations.

5 Accounting for the Effect of the SARS Shock

SARS reduced imports, more strongly for less diversified firms. The question that remains unanswered is how much the SARS shock on imports had raised firms' marginal costs and reduced aggregate output. The lack of domestic sourcing data prevents me from doing a full-fledged structural estimation to uncover all underlying parameters following AFT. Furthermore, I have only examined the response in the intensive margin. Despite these, the following proposition shows that answering this question requires only estimating the demand elasticity, observing the pre-shock sourcing behavior, and the estimated effects on imports in the intensive margin.

Proposition 7. *Under input complementarity, the change in firms' marginal cost in the face of adverse shocks to inputs can be inferred as:*

$$\widehat{c}_i(\varphi) = (\sum_{j\times k\in\mathcal{C}(\varphi)} \chi_{ijk}(\varphi)\widehat{M}_{ijk}(\varphi))^{\frac{1}{1-\sigma}},$$

with the pre-shock sourcing intensities $\{\chi_{ijk}(\varphi)\}$, the estimated change in trade flows $\{\widehat{M}_{ijk}(\varphi)\}$ and the demand elasticity σ.

Proof. See Appendix 8.7 □

Although marginal cost is not directly observable and cannot be easily estimated. This result tells us that $\{\chi_{ijk}(\varphi)\}$, $\{\widehat{M}_{ijk}(\varphi)\}$ and σ are *sufficient statistics* to inferred its change due to the shock. $\{\chi_{ijk}(\varphi)\}$ is available from the data. $\{\widehat{M}_{ijk}(\varphi)\}$ can be calculated given the estimates from the previous section. The only unknown is the demand elasticity σ. As I have estimated, $\sigma - 1 - \theta = 0.464$ from the interaction term of Column (4) in Table 1 which corresponds to coefficient β in Equation (4.6). According to Proposition 5, β's theoretical counterpart is $-(\sigma - 1 - \theta)$. If θ is estimated, σ can also be inferred. I next estimate θ by exploring firms' sourcing decision with respect to tariff variations across markets.

5.1 Estimating Efficiency Dispersion Parameter θ

The key relationship that I use to estimate θ is $\chi_{ijk}(\varphi) = \frac{A_j T_k(\tau_{ij}\tau_{jk}w_j w_k)^{-\theta}}{\Psi_i(\varphi)}$, which is the sourcing intensity from origin k through customs district j for a firm with productivity φ in region i.

Suppose $\chi_i^d(\varphi) \equiv \frac{\phi_i^d(\varphi)}{\Psi_i(\varphi)}$ is the intensity of domestic sourcing where $\phi_i^d(\varphi)$ is the capability of domestic sourcing, a ratio-type estimator can be formulated:

$$\begin{aligned}
\ln \chi_{ijk}(\varphi) - \ln \chi_i^d(\varphi) &= \ln \frac{A_j T_k (\tau_{ij}\tau_{jk} w_k w_j)^{-\theta}}{\Psi_i(\varphi)} - \ln \frac{\phi_i^d(\varphi)}{\Psi_i(\varphi)} \\
&= \underbrace{\ln A_j {w_j}^{-\theta}}_{customs\ FE} + \underbrace{\ln T_k {w_k}^{-\theta}}_{origin\ \mathrm{FE}} - \theta \ln \tau_{ij} - \theta \ln \tau_{jk} - \ln \phi_i^d(\varphi)
\end{aligned} \tag{5.7}$$

AFT normalize $\phi_i^d(\varphi)$ to be the value one which implies that all importers have the same domestic sourcing capability. Hence $\ln \phi_i^d(\varphi) = 0$ and disappears from the equation above. Since I observe the locations of firms and other firm variables, I allow it to vary cross firms by controlling firm characteristics and firm fixed effect.[41]

Compared with AFT, I also allow for domestic trade costs τ_{ij} and intermediation efficiencies at customs district $A_j {w_j}^{-\theta}$ to affect firms sourcing behaviour. In the equation above, I control for $\ln A_j {w_j}^{-\theta}$ and $\ln T_k {w_k}^{-\theta}$ by customs fixed effect and origin fixed effect, respectively. For domestic and international trade costs, I assume that $\ln \tau_{ij}^{\theta} = \alpha_0 + \alpha_1 \ln dist_{ij} + \alpha_2 comLang_{ij} + \alpha_3 comCustoms_{ij} + \epsilon_{ij}$ and $\ln \tau_{jk}^{\theta} = \beta_0 + \beta_1 \ln dist_{jk} + \beta_2 coCHN_{jk} + \beta_3 t_k + \varepsilon_{jk}$. Domestic distances $dist_{ij}$ are measured in great circle distance between the prefecture where the firm is located and the importing customs district. The coordinates of the Chinese prefectures are measured in ArcGIS as the centroid of each prefecture. The coordinates of the customs districts are measured as the centroid of the major gateway city within the customs district.[42] Distances between Chinese customs districts and sourcing origins $dist_{jk}$ are measured in terms of great circle distances between the centroids of major gateways as well.[43] $comLang_{ij}$ is a dummy variable indicating whether the domestic destination i shares the same language as the customs district j. This variable is coded using the Language Atlas of China in Lavely (2000) which provides data at county level. I further aggregate the data to prefecture city and customs district level. $comCustoms_{ij}$ is a dummy indicating whether destination i is within the customs district j or not. It is meant to capture the trade costs imposed by the customs administrative

[41]Firms with high domestic sourcing capability are very likely to have high global sourcing capability. The global sourcing capability will be overestimated unless $\phi_i^d(\varphi)$ is controlled. Yet firms in regions with poor access to foreign markets may source more from the domestic market. In a different context, Baum-Snow *et al* (2016) show that domestic and foreign market access have different implications for Chinese urban growth.

[42]When there is more than one major gateway city, the minimum of the distances from the ports is used. The list of major gateways for each customs district is in Appendix A10.

[43]For coastal countries, I identify the largest port. For inland countries, the capital city is used. I also seek a robustness check with maritime distance. It is computed as using a map of maritime shipping from Halpern et al. (2015) to extract the shortest path connecting the ports. I then calculate the length of the path. The result is quantitatively similar.

boundaries. Next, since Rauch and Trindade (2002) find that ethnic Chinese networks facilitate trade between countries, I construct the variable $coCHN_{jk}$ which is the share of ethnic Chinese in origin k multiplied by the share of overseas Chinese for customs district j. Historically, some Chinese regions such as Guangdong had more emigrants bound for other countries. These regions may have formed a closer network of foreign suppliers and enjoyed lower trade costs than other regions in China. The share of ethnic Chinese for each origin is from Poston Jr et al. (1994).[44] The share of overseas Chinese for customs districts is constructed using the Chinese City Yearbook 1995. The Yearbook reports the number of overseas Chinese for each prefecture.[45] I aggregate it up to customs district level and divide it by the local population. The result is reported in the last column of Appendix Table A10. Finally, I construct firm-market import tariffs using data from TRAINS following Fitzgerald and Haller (2014). The tariff is constructed as $t_k^n = \sum_p \frac{ps_t^n + ps_{t-1}^n}{2} ln(1 + t_{k,p})$ where ps_t^n is the product share in firm n's import basket at period t and $t_{k,p}$ is the import tariff imposed by China on product p from origin k. It varies by market since product tariffs vary by market. Such variations shift the costs of sourcing and allow us to identify the dispersion parameter θ.

In the end, the equation that I estimate is:

$$\begin{aligned}\ln \chi_{ijk}^n - \ln \chi_i^{dn} \quad &= \quad a + C_j + O_k - \alpha_1 \ln dist_{ij} - \alpha_2 comLang_{ij} - \alpha_3 comCustoms_{ij} \\ &\quad -\beta_1 \ln dist_{jk} - \beta_2 coCHN_{jk} - \beta_3 ln(t_k^n) + X_i^n \delta + F^n + \xi_{ijk}^n,\end{aligned}$$

where C_j and O_k are custom area and origin fixed effects respectively. Firm characteristics X_i^n and firm fixed effect F^n capture the unobserved domestic sourcing capability $\ln \phi_i^d(\varphi)$ in Equation (5.7). β_3 is the coefficient of interest which corresponds to θ. The results for the estimation using year 2006 data are reported in Table 5.[46] The main specification of interest is shown in Column (4) where we have $\widehat{\theta} = 5.50$. This is close to the value in the literature as surveyed by Head and Mayer (2014). They find a median trade elasticity at 5.03 for structural estimations using tariff variations.[47] To address the concern that current product share is endogenous to current tariff in Fitzgerald and Haller (2014), I also use a tariff measure which

[44]The sample used by Rauch and Trindade (2002) is limited to a smaller number of countries for which the gravity variables are available. I use the full sample from Poston Jr et al. (1994).

[45]In the data they are called "Hua2qiao2" in Pinyin, which means 'overseas Chinese'.

[46]There was not much variation in import tariffs across markets for China before 2006. Most Chinese trade agreements took effect after 2005. Before that, China imposed homogeneous import tariffs across markets except several least developed African countries.

[47]This is also their preferred value. AFT estimate θ using the variation in wages across sourcing origins and get a much lower elasticity at 1.789.

only use the lagged product shares as weights. Instead of controlling gravity variables, I use origin-customs district-destination fixed effect to fully absorb the iceberg trade costs. I also try other specifications and the elasticity remains robust as in the Appendix Table A11. Given the estimate that $\widehat{\theta} = 5.50$, the demand elasticity is $\widehat{\sigma} = 5.50 + 1 + 0.465 = 6.965$.

The estimated origin fixed effect estimated corresponds to $\ln T_k w_k^{-\theta}$ which capture the efficiency of each origin. I plot it against total imports from each origin in Appendix Figure A1 (a). The estimated customs district fixed effect captures the efficiency of each customs $\ln A_j w_j^{-\theta}$. Panel (b) plots it against total imports through each customs. The estimated efficiency of Shenzhen is about 1.9 times higher than Shanghai. This probably partly explains why imports through Shenzhen were almost the same as Shanghai but the number of importers imported via Shenzhen was just about 1/3 of Shanghai.

5.2 Effect of SARS on Firms' Marginal Cost and Aggregate Output

With the estimated demand elasticity σ, we can now estimate the effect on firms' marginal cost using Proposition 7. Using the estimated effect of SARS on imports from the previous section, I compute the point estimate of changes in imports by $\widehat{M}_{ijk}(\varphi) = \tilde{\alpha}_2 SARS_{jk,t} + \tilde{\beta}\chi_{ijk} SARS_{jk,t} + \tilde{\gamma} CoSARS_{ijk}^{nt}$ with $\tilde{\alpha}_2 = -0.0555$, $\tilde{\beta} = -0.464$, and $\tilde{\gamma} = 0.00454$ according to column (4) of Table 1. Figure 6 (a) plots the estimated changes in marginal cost against the number of sourcing routes for the affected firms.[48] The average effect on marginal cost is about 0.7%. Interestingly, there is a general downward sloping trend: firms sourcing from more routes appear less affected. However, if the homogeneous pass-through specification is used to compute $\widehat{M}_{ijk}(\varphi) = \tilde{\alpha}_2 SARS_{jk,t}$ where $\tilde{\alpha}_2 = -0.0794$ according to column (2) of Table 1, the result plotted in panel (b) shows an opposite trend: the more diversified firms appear less resilient. This highlights how heterogeneous pass-through is translated to heterogeneity in resilience.

With the estimated effect on marginal cost, the effect on firm revenue is:

$$\widehat{R}_i(\varphi) = \widehat{c}_i(\varphi)^{1-\sigma}. \tag{5.8}$$

Therefore, given that $\sigma = 6.965$, an 1% increase in marginal cost translates to a loss of $(1-\sigma)\% = -5.965\%$ in revenues.[49] Before showing the aggregate effect, I examine whether these firm-level

[48] The figure is generated using local polynomial regression, dropping the top 1% firms in terms of the number of sourcing routes used.

[49] To the extent that σ might be different across firms, it has been investigated by Yeh (2016). He found that larger firms face smaller price elasticities. This is an additional channel that they might respond less to shocks.

Table 5: Estimating efficiency dispersion parameter θ

Dependent Variable: foreign sourcing relative to home sourcing $ln(\chi_{ijk}) - ln\frac{\phi^d}{\Phi}$	(1)	(2)	(3)	(4)
ln geodist customs district-destination	-0.166***	-0.0977***	-0.0521***	-0.0522***
	(0.0125)	(0.0156)	(0.0159)	(0.0159)
ln geodist origin-customs district	-0.437***	-0.435***	-0.408***	-0.411***
	(0.0141)	(0.0140)	(0.0153)	(0.0153)
common customs district		0.473***	0.446***	0.446***
		(0.0682)	(0.0673)	(0.0673)
common language customs district-destination			0.842***	0.845***
			(0.0880)	(0.0880)
co-Chinese			10.55***	10.48***
			(2.630)	(2.627)
FH firm-market import tariff				-5.500***
				(0.797)
Firm, Industry, Ownership, Origin, Customs, Region FE	Y	Y	Y	Y
R^2	0.456	0.457	0.458	0.458
No. of observations	121742	121742	121742	121742

Notes: The dependent variable is the the log difference of probability in sourcing from a route relative to sourcing at home. The sample only includes importers in year 2006 that are not entrants in year 2005. "co-Chinese" is the share of ethnic Chinese in the origin multiplied by the share overseas Chinese in the Chinese customs district. "FH firm-market import tariff" a firm market specific tariff constructed following Fitzgerald and Haller (2014). It is a weighted average of product tariffs using the basket goods in current and lagged years. The numbers in the parentheses are standard error clustered at firm level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

revenue shocks are meaningful. I regress the actual firm revenue growth rates in year 2003 on the accumulated revenue shocks over the quarters that the firm was affected in 2003.[50] The result is shown in Table 6. Columns (1) to (3) include both importers and non-importers. Zeros are assigned for the accumulated shocks for non-importers. Columns (4) and (5) only include importers. As we can see, in all regressions, firms which had larger revenue shocks due to SARS had a lower growth rate in 2003. So the constructed shocks are indeed correlated with the actual growth rates.

I then aggregate the loss in revenue across firms using

$$\widehat{R} = \sum \frac{R_i(\varphi)}{R} \widehat{c}_i(\varphi)^{1-\sigma}$$

to infer the total output loss where $\frac{R_i(\varphi)}{R}$ is the observed output share of firm i from the data. I do this exercise for each quarter that SARS was affecting the economy. The result is plotted in Figure 7. At the peak of the epidemic 2003Q2, SARS led to a loss of about 0.7% in Chinese manufacturing output. It quickly subsided to just 0.2% when the epidemic ended in 2003Q3.[51]

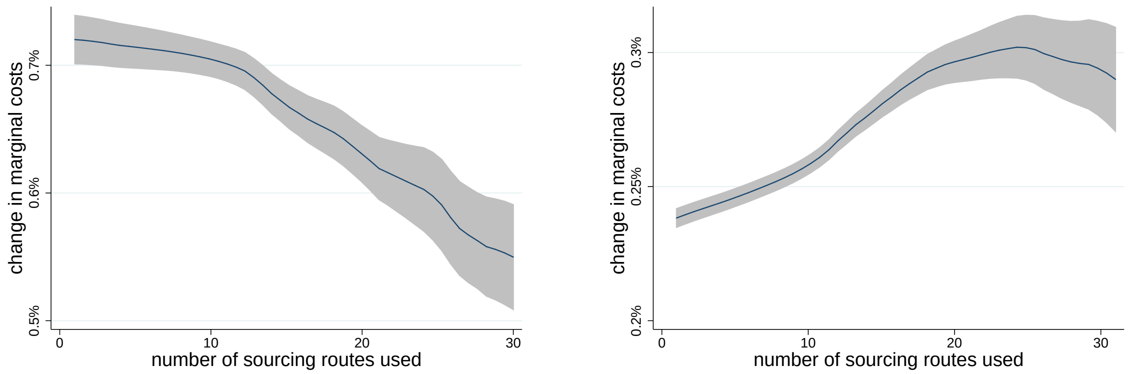
Table 6: Verifying the firm level revenue shock

Dependent Variable:	All Firms			Importers Only	
firm revenue growth rate in year 2003	(1)	(2)	(3)	(4)	(5)
accumulated revenue loss due to SARS shocks on imports	-0.881***	-0.705***	-0.770***	-0.569***	-0.410***
	(0.207)	(0.204)	(0.214)	(0.119)	(0.109)
ln firm age		-0.319***	-0.344***		-0.535***
		(0.0445)	(0.0472)		(0.0787)
ln firm employment			0.100***		0.0415**
			(0.0305)		(0.0197)
Prefecture FE	Y	Y	Y	Y	Y
industry FE	Y	Y	Y	Y	Y
Ownership FE	Y	Y	Y	Y	Y
R^2	0.00570	0.00614	0.00623	0.0573	0.0724
No. of observations	140081	140081	140081	11585	11585

Notes: The dependent variable is the the growth rate of firm revenue in 2003. Columns (1) to (3) include both importers and non-importers. Columns (4) and (5) only include importers. The numbers in the parentheses are standard error clustered at industry-prefecture level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

[50] Quarterly level firm outputs are not observable since CAIS reports firm revenues at annual frequency. I sum the inferred revenue shocks over the quarters that firms were affected. For example, if a firm had a revenue shock of 1% in Spring and 2% in Summer, the overall shock is 3%.

[51] Lee and McKibbin (2004) simulated a CGE model to estimate the economic impact of SARS. They find a reduction of Chinese GDP by 0.37%. Since manufacturing is about 32.5% of Chinese GDP in 2003, my estimation implies that GDP fell by 32.5%*0.7%=0.23% at the peak of SARS due to shocks on imports.



(a) Heterogeneous Pass-through

(b) Homogeneous Pass-through

Figure 6: Effect on firms' marginal costs

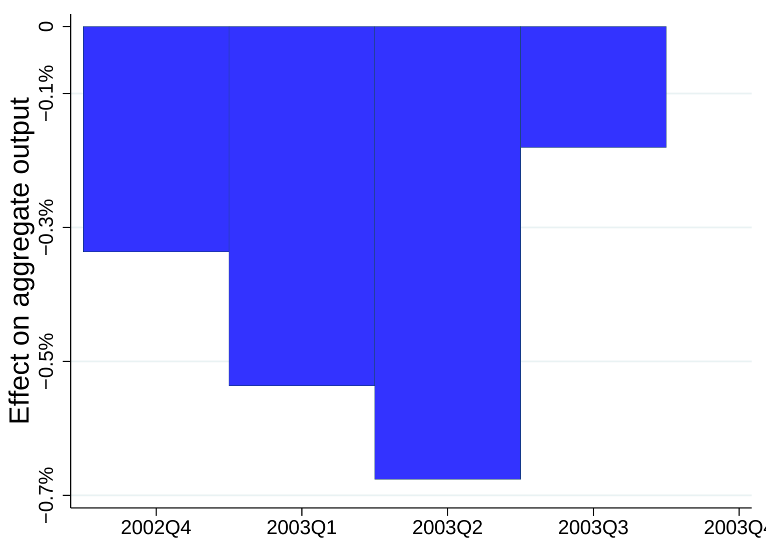


Figure 7: Effect on aggregate output

6 Diversification, Roads and Resilience

I have provided evidence showing that sourcing diversification made firms more resilient to the SARS epidemic. But the questions remained are: (a) who are more diversified, and (b) can we improve firms' resilience by making them more diversified in sourcing? I will address these two questions in this section. Answering these two questions not only provides further tests on the model, but also helps us to identify barriers that keep firms from being resilient and find policies to improve their resilience.

6.1 Productivity and Diversification

Proposition 2 predicts diversification depends on productivity: high productivity firms are more diversified as measured by the HHI of their sourcing diversification. To test this prediction, I run the following regression:

$$HHI_{nt} = a_0 + a_1 lnF_{nt} + \sum_k \beta_k X_k^n + \epsilon_{nt},$$

where HHI_{nt} is the HHI of firm n at period t, F_{nt} is the firm productivity, and X_k is the other firm characteristics. a_1 is the main coefficient of interest. According to the proposition, we should expect $a_1 < 0$. As mentioned before, HHI is constructed as the sum over the squares of input expenditure share in each trade route for each firm.[52] It is assigned as one for non-importers when I look at the full sample of firms since I do not observe domestic sourcing.

The results are reported in Table 7. Columns (1) and (2) use the full sample, including importers and non-importers. Columns (3) and (4) only look at importers. The controls include year, ownership, industry, and region fixed effects in all columns, and firm fixed effects in columns (2) and (4). Across all columns, we find that the estimated $\widehat{a}_1$ is negative and highly significant. So indeed, consistent with the model, high productivity firms are more diversified in sourcing. These empirical findings also implies that it is important to control for firm productivity when examining sourcing diversification.

6.2 Roads and Diversification

In this subsection, I examine the role of infrastructure, specifically railways and highways which have been shown to reduce trade costs (e.g., Donaldson, forthcoming), in shaping firms' sourcing diversifications and resilience to SARS. If infrastructure helps firms to become more diversified in sourcing, it can also make them to more resilient to SARS given the results from section 4.

According to Proposition 3, firms' sourcing strategies become more diversified along the extensive margin if there is reduction in trade costs. To test this proposition, I explore the expansion of the Chinese highway and railway network from 2000-2006. I examine whether

[52]Following AFT, firms with imports more than its total inputs are excluded from the sample. Imports on fuels and mineral products are not counted. Wage bills are included as total inputs to address the concern on home sourcing.

Table 7: Firm Productivity and Diversification of Sourcing: all firms

Dependent Variable:	All Firms		Importers Only	
sourcing diversification in HHI	(1)	(2)	(3)	(4)
ln TFP	-0.00839***	-0.000853***	-0.0835***	-0.0223***
	(0.000168)	(0.000118)	(0.00113)	(0.00156)
Year FE	Y	Y	Y	Y
Ownership FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Region FE	Y	Y	Y	Y
Firm FE	N	Y	N	Y
R^2	0.367	0.860	0.160	0.597
No. of observations	1328727	1224458	185957	165863

Notes: The numbers in the parentheses are standard error clustered at firm level. The number of observations varies across regressions as I use the *reghdfe* command in Stata which gets rid of singletons for fixed effects nested with each other. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

firms in regions connected to highways or railways expand their sourcing strategies or not by running the following regression:

$$ln(1 + N_{it}^n) = b_0 + b_1 Highway_{it} + b_2 Railway_{it} + \delta_i + \delta_n + \delta_t + \sum_k \beta_k X_k^{nt} + \zeta_{it}^n,$$

N_{it}^n counts the extensive margin of the sourcing strategy for firm n at period t. $Highway_{it}$ and $Railway_{it}$ are dummies indicating region i's connection to highways and railways, respectively. δ_i, δ_n and δ_t capture region, firm, and time fixed effects, respectively. X_k^{nt} is added to control various firm characteristics, including firm productivity. ζ_{it}^n is the error term.

To implement this regression, I construct dummies indicating whether or not a region was connected to highways or railways using China GIS data provided by the ACASIAN Data Center at Griffith University. The data provide year 1999 county boundaries and transportation data for the years 2000, 2004 and 2007. To construct the highway and railway network for each year between 2000 and 2006, I manually collect the opening date of Chinese highway and railway by section according to news reports, government reports, and other online sources that I collect.[53] The complete networks for years 2000 and 2006 are illustrated in Figure A3 and A4. As we can see, the railway network was already quite dense in 2000.[54] The highway network was mostly confined to the coastal provinces in 2000, but expanded quickly to most of the country in 2006. The geographical unit of my analysis is a county.[55] I use the region code in CAIS to identify

[53]If there is a conflict with ACASIAN data on the opening date, I follow the sources that I have collected.

[54]The recent impressive development of high speed railways was mostly part of the stimulation package after the 2008 financial crisis.

[55]The base map from ACASIAN combines the urban districts into a single urban unit for each prefecture. I

the county that each firm is located.[56]

The results presented in Table 8 include counts of customs districts, origins, and customs district-country pairs as outcome variables. Columns (1) to (3) use the full sample of importers. A well-known issue in the literature on infrastructure evaluation is the endogenous placement of roads. If roads were built to connect importers, the estimated effect would be upward biased. To handle this issue, I follow the "inconsequential unit approach" (Chandra and Thompson, 2000) to exclude firms located on the end nodes of the network.[57] The idea is that the unobserved characteristics of the units between the nodes of the network should be inconsequential to the placement the roads. These units got connected simply because they lie between the nodes. Thus I exclude importers located in urban units within each city and provincial capital cities, since highways or railways were built to connect these regions. The results are presented in columns (4) to (6). As is obvious across all the columns, connection to the highway increases the number of customs districts, sourcing origins, and customs district-origin pairs in importers' sourcing strategy. The effect is significant and robust across different samples and outcomes. However, connection to railways does not appear to have a significant effect.

Although Proposition 3 only makes prediction about diversification along the extensive margin, firms are likely be more diversified as measured by HHI if they have a wider sourcing strategy. I expect that firms got connected by highways or railways will have a lower HHI. This is formally tested by regressing HHI on dummies of highway and railway connections. The results are presented in Table 9. Connections with railways and highways are associated with more diversification, as indicated in columns (1) to (3). In column (4), I show that the same result holds after excluding the firms from the urban units and provincial capitals. Although railways do not appear to have a significant effect on the extensive margin, they help firms to diversify when the intensive margin is taken into account.

6.3 Roads and Resilience

I have shown that diversification makes firms more resilient to the SARS epidemic and roads help to increase diversification. The remaining question is: do roads increase firms' resilience to the SARS epidemic? The idea is that if firms in regions with railway or highway connection are more diversified, this should make them more resilient. To see if this is the case, I run the

include these urban units in my baseline result and exclude them in the robustness checks.

[56]I use the region code (in Chinese) from the Ministry of Civil Affairs to link region codes over time.

[57]Redding and Turner (2015) synthesise the literature that addresses the endogenous placement of roads.

Table 8: Roads and sourcing diversification: the extensive margin

	Full sample			Excluding nodes of road network		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: $ln(N+1)$ in	customs districts	origins	customs districts-origins	customs districts	origins	customs districts-origins
highway connected	0.00531***	0.00514***	0.00499***	0.00414***	0.00444***	0.00414**
	(0.000938)	(0.00142)	(0.00150)	(0.00102)	(0.00157)	(0.00165)
rail connected	-0.00105	0.00146	0.000790	-0.00169	0.00179	0.00109
	(0.00137)	(0.00220)	(0.00229)	(0.00154)	(0.00239)	(0.00250)
ln TFP	0.00771***	0.0159***	0.0168***	0.00717***	0.0144***	0.0151***
	(0.000280)	(0.000468)	(0.000491)	(0.000329)	(0.000557)	(0.000580)
ln age	0.00949***	0.0162***	0.0170***	0.00815***	0.0136***	0.0141***
	(0.000457)	(0.000732)	(0.000773)	(0.000493)	(0.000797)	(0.000832)
Year FE	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
Ownership FE	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y
region FE	Y	Y	Y	Y	Y	Y
R^2	0.811	0.873	0.871	0.806	0.873	0.872
No. of observations	1223731	1223731	1223731	832414	832414	832414

Notes: Columns (1) to (3) use the full sample while columns (4) to (6) exclude firms located in urban units and provincial capitals. The numbers in the parentheses are standard error generated from observed information matrix. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table 9: Roads and sourcing diversification: HHI

Dependent variable: HHI	Full sample			Excluding nodes of road network
	(1)	(2)	(3)	(4)
connected to highway	-0.00310***		-0.00305***	-0.00258***
	(0.000315)		(0.000315)	(0.000354)
connected to railway		-0.00168***	-0.00139***	-0.00249***
		(0.000490)	(0.000490)	(0.000516)
ln TFP	-0.000831***	-0.000826***	-0.000827***	-0.000417***
	(0.000118)	(0.000118)	(0.000118)	(0.000145)
ln firm age	-0.000716***	-0.000717***	-0.000720***	-0.000768***
	(0.000144)	(0.000144)	(0.000144)	(0.000163)
Year FE	Y	Y	Y	Y
Ownership FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Region FE	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
R^2	0.860	0.860	0.860	0.865
No. of observations	1223731	1223731	1223731	832414

Notes: The dependent variable is firms' concentration of sourcing measured by the Herfindahl-Hirschman Index (HHI). Columns (1) to (3) use the full sample while column (4) excludes importers located in urban units and provincial capitals which are the nodes of the road network. The numbers in the parentheses are standard error clustered at firm level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

following regression:

$$\ln Import^{nt}_{ijk} = c_0 + c_1 Highway_{it} + c_2 Railway_{it} + \gamma_0 SARS_{jk,t}$$
$$+\gamma_1 Highway_{it} SARS_{jk,t} + \gamma_2 Railway_{it} SARS_{jk,t} + \sum \beta_k X^n_{kt} + \epsilon_{nt},$$

where $Highway_{it}$ and $Railway_{it}$ are the connection dummies and $SARS_{jk,t}$ is the dummy indicating whether route jk was affected by SARS or not at period t. The key coefficients of interest are γ_1 and γ_2. If connectivity to roads indeed increases resilience, we expect them to be positive. The results are presented in Table 10. One difference from the previous section is that I cannot control for the firm-time fixed effect because the connectivity dummies $Highway_{it}$ and $Railway_{it}$ are defined at region-time level. Firm-time fixed effects or region-time fixed effects are fully multi-collinear with these dummies. Instead, I control for province-time fixed effects to handle demand or productivity shocks due to SARS.

Table 10: Roads and Resilience

Dependent Variable:	Full sample		Excluding nodes of road network	
firm import by route $ln(imp_{ijk})$	(1)	(2)	(3)	(4)
trade route hit by SARS=1	-0.0924***	-0.168***	-0.121***	-0.198***
	(0.0195)	(0.0475)	(0.0244)	(0.0547)
highway connected=1		0.0332		0.000513
		(0.0337)		(0.0392)
trade route hit by SARS=1 x highway connected=1		0.0476		0.0382
		(0.0388)		(0.0434)
railway connected=1		0.0414		0.133**
		(0.0357)		(0.0622)
trade route hit by SARS=1 x railway connected=1		0.0432		0.0637*
		(0.0307)		(0.0350)
province-time, Destination, Origin, Customs, Ownership, Industry FE	Y	Y	Y	Y
R^2	0.111	0.111	0.123	0.123
No. of observations	2231882	2231882	1348241	1348241

Notes: A firm is defined as connected to highway if the county that the firm located is connected to the highway in each year. Columns (1) to (2) include full sample while columns (3) and (4) exclude firms located in urban units or provincial capitals. The numbers in the parentheses are standard error clustered at industry and prefecture level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

The results are shown in Table 10. Column (1) shows the average effect of the SARS shock. Columns (2) study the connectivity to highways and railways. Both highways and railways appear to dampen the effect of the SARS shock but the effect is not significant. When I exclude firms located in urban units and provincial capitals in column (4). The dampening effect of railways appears to be larger and marginally significant.[58] Overall, connection to railways reduced the negative impact of SARS on imports by about 6%. However, the effect of highway connectivity remains insignificant.

7 Conclusion

This paper studies how diversification shapes the resilience to adverse shock along global supply chains. The Chinese firms which are more geographically diversified in their input sourcing appear less volatile. However, geographical diversification is not a free lunch. Gravity drags firms to source through closer customs districts. Only the larger and more productive firms manage to source through more customs districts. I build a model to account for these facts based on the work by Antràs, Fort and Tintelnot (2017). The model predicts that high productivity firms are more diversified, hence more resilient to adverse shocks under input complementarity. It also predicts that trade liberalization or improvement in infrastructure facilitates diversification. I explore the 2003 SARS epidemic as a natural experiment to test the model predictions and find that, the damage on imports is indeed smaller if the firm is more diversified hence need not rely so much on sourcing routes hit by SARS. Moreover, connection with highways and railways appears to facilitate diversification in sourcing and reduce the impact of the SARS shock. This is a benefit of infrastructure that should not be overlooked by policy makers.

References

Allen, T. and Atkin, D., 2015. Volatility, Insurance, and the Gains from Trade. Working paper.

Allen, T. and Arkolakis, C., 2014. Trade and the Topography of the Spatial Economy. The Quarterly Journal of Economics, 129(3), pp.1085-1140.

Amiti, M., Itskhoki, O. and Konings, J., 2014. Importers, Exporters, and Exchange Rate Disconnect. The American Economic Review, 104(7), pp.1942-1978.

[58] To contain the spread of SARS, local Chinese governments set up check-points on highways to examine the temperature of drivers and passengers. While such checks were also applied to passengers travelling by railway before boarding, they were unlikely to disrupt trains, which follow fixed schedules, especially those only carrying goods. In contrast, so many check-points were set up on roads that the Ministry of Public Security had to issue an executive order called "Five Forbidden Things"(in Chinese) in May 2003: no interruptions of traffic were allowed in the name of fighting against SARS; no traffic controls at provincial borders; no road blocks to stop traffic; no forced U-turns for vehicles in the normal course of things; and no traffic jams in the name of quarantine. Such a difference probably explains the different effect of railways and highways connection.

Antràs, P. and de Gortari, A., 2017. On the Geography of Global Value Chains. Working paper No.23456. National Bureau of Economic Research.

Antràs, Pol, Teresa C. Fort, and Felix Tintelnot. 2017. The Margins of Global Sourcing: Theory and Evidence from US Firms. American Economic Review 107 (9).

Antras, P. and Helpman, E., 2004. Global Sourcing. Journal of Political Economy, 112(3), pp.552-580.

Atkin, D. and Donaldson, D., 2015. Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs. No. 21439. National Bureau of Economic Research.

Barrot, J.N. and Sauvagnat, J., 2016. Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks. The Quarterly Journal of Economics, 131(3), pp.1543-1592.

Baum-Snow, Nathaniel, Loren Brandt, J. Vernon Henderson, Matthew A. Turner, and Qinghua Zhang., 2016. Highways, Market Access, and Urban Growth in China. Spatial Economics Research Centre, LSE.

Berman, N., Martin, P. and Mayer, T., 2012. How Do Different Exporters React to Exchange Rate Changes?. The Quarterly Journal of Economics, 127(1), pp.437-492.

Bernard, A.B., Jensen, J.B., Redding, S.J. and Schott, P.K., 2007. Firms in International Trade. The Journal of Economic Perspectives, 21(3), pp.105-130.

Bernard, Andrew B., Andreas Moxnes, and Yukiko U. Saito., 2015. Production Networks, Geography and Firm Performance. No.21082. National Bureau of Economic Research.

Blaum, Joaquin, Claire Lelarge, and Michael Peters. 2016. The Gains From Input Trade with Heterogeneous Importers. Working paper, Brown University.

Boehm, C., Flaaen, A. and Pandalai-Nayar, N., 2015. Input Linkages and the Transmission of Shocks: Firm-Level Evidence from the 2011 Tōhoku Earthquake. Working paper.

Brandt, L., J. Van Biesebroeck, and Y. Zhang. Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing. Journal of Development Economics 97, No. 2 (2012): 339-351.

Burgess, R. and Donaldson, D., 2010. Can Openness Mitigate the Effects of Weather Shocks? Evidence from India's Famine Era. The American Economic Review, 100(2), pp.449-453.

Burgess, R. and Donaldson, D., 2012. Railroads and the Demise of Famine in Colonial India. Working Paper.

Carvalho, V.M., Nirei, M., Saito, Y.U. and Tahbaz-Salehi, A., 2016. Supply Chain Disruptions: Evidence from the Great East Japan Earthquake. Working paper.

Caselli, F., Koren, M., Lisicky, M. and Tenreyro, S., 2015. Diversification through Trade. No. 21498. National Bureau of Economic Research.

Chandra, A. and Thompson, E., 2000. Does Public Infrastructure Affect Economic Activity?: Evidence from the Rural Interstate Highway System. Regional Science and Urban Economics, 30(4), pp.457-490.

Christopher, M. and Peck, H., 2004. Building the Resilient Supply Chain. The International Journal of Logistics Management, 15(2), pp.1-14.

Correia, S., 2016. REGHDFE: Stata Module to Perform Linear or Instrumental-Variable Regression Absorbing Any Number of High-Dimensional Fixed Effects. Statistical Software Components.

Coşar, A. Kerem, and Banu Demir. Domestic Road Infrastructure and International Trade: Evidence from Turkey. Journal of Development Economics 118 (2016): 232-244.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, 2007. Unbalanced Trade, American Economic Review, vol. 97(2), pages 351-355, May.

Donaldson, Dave. Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. American Economic Review, forthcoming.

Eaton, J. and Kortum, S., 2002. Technology, Geography, and Trade. Econometrica, 70(5), pp.1741-1779.

Esposito, F., 2017. Entrepreneurial Risk and Diversification through Trade. Working Paper, Tufts University.

Fajgelbaum, P. and Redding, S.J., 2014. External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914 (No. w20217). National Bureau of Economic Research.

FAO, 2016. Impact of the Ebola Virus Disease Outbreak on Market chains and Trade of Agricultural Products in West Africa. Policy report.

Fillat, J.L. and Garetto, S., 2015. Risk, Returns, and Multinational Production. The Quarterly Journal of Economics, 130(4), pp.2027-2073.

Fitzgerald, D. and Haller, S., 2014. Exporters and Shocks: Dissecting the International Elasticity Puzzle. No. 19968. National Bureau of Economic Research.

Feenstra, R.C., 1994. New Product Varieties and the Measurement of International Prices. The American Economic Review, pp.157-177.

Feenstra, R.C. and Hanson, G.H., 2005. Ownership and Control in Outsourcing to China: Estimating the Property-Rights Theory of the Firm. The Quarterly Journal of Economics, 120(2), pp.729-761.

Furusawa, Taiji, Tomohiko Inui, Keiko Ito, and Heiwai Tang. Offshoring, Relationship-Specificity, and Domestic Production Networks. Discussion Paper 15122, Research Institute of Economy, Trade and Industry (RIETI), 2015.

Goldberg, Pinelopi K., Amit Khandelwal, Nina Pavcnik, and Petia Topalova. Imported Intermediate Inputs and Domestic Product Growth: Evidence from India. Quarterly Journal of Economics, 125(4), 2010, pp. 1727-67.

Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., Lowndes, J.S., Rockwood, R.C., Selig, E.R., Selkoe, K.A. and Walbridge, S.,2015. Spatial and Temporal Changes in Cumulative Human Impacts on the World's Ocean. Nature Communications.

Halpern, L., Koren, M. and Szeidl, A., 2015. Imported Inputs and Productivity. The American Economic Review, 105(12), pp.3660-3703.

Head, K., Jing, R. and Ries, J., 2017. Import Sourcing of Chinese Cities: Order versus Randomness. Journal of International Economics, 105, pp.119-129.

Head, K. and Mayer, T., 2014. Gravity Equations: Workhorse, Toolkit, and Cookbook, in the Handbook of International Economics Vol. 4, eds. Gopinath, Helpman, and Rogoff.

Heymann, D.L., Mackenzie, J.S. and Peiris, M., 2013. SARS Legacy: Outbreak Reporting is Expected and Respected. The Lancet, 381(9869), pp.779-781.

Hollingsworth, T.D., Ferguson, N.M. and Anderson, R.M., 2006. Will Travel Restrictions Control the International Spread of Pandemic Influenza?. Nature Medicine, 12(5), pp.497-499.

Hummels, D., Ishii, J. and Yi, K.M., 2001. The Nature and Growth of Vertical Specialization in World Trade. Journal of International Economics, 54(1), pp.75-96.

Jones, R.W., 1965. The Structure of Simple General Equilibrium Models. Journal of Political Economy, 73(6), pp.557-572.

Kramarz, F., Martin, J. and Mejean, I., 2016. Volatility in the Small and in the Large: The Lack of Diversification in International Trade. Working paper.

Koren, M. and Tenreyro, S., 2013. Technological Diversification. The American Economic Review, 103(1), pp.378-414.

Kugler, Maurice, and Eric Verhoogen. Plants and Imported Inputs: New Facts and an Interpretation. The American Economic Review 99, no. 2 (2009): 501-507.

Kurz, C. and Senses, M.Z., 2016. Importing, Exporting, and Firm-Level Employment Volatility. Journal of International Economics, 98, pp.160-175.

Lavely, William. Coding Scheme for the Language Atlas of China. University of Washington. Memo. 2000

Levinsohn, James, and Amil Petrin.Estimating Production Functions Using Inputs to Control for Unobservables. The Review of Economic Studies 70, no. 2 (2003): 317-341.

Lee, J.W. and McKibbin, W.J., 2004. Globalization and Disease: the Case of SARS. Asian Economic Papers, 3(1), pp.113-131.

Melitz, Marc J. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Econometrica 71, no. 6 (2003): 1695-1725.

Manova, K. and Yu, Z., 2016. How Firms Export: Processing Vs. Ordinary Trade with Financial Frictions. Journal of International Economics, 100, pp.120-137.

Ponomarov, S.Y. and Holcomb, M.C., 2009. Understanding the Concept of Supply Chain Resilience. The International Journal of Logistics Management, 20(1), pp.124-143.

Poston Jr, Dudley L., Michael Xinxiang Mao, and Mei-Yu Yu. The Global Distribution of the Overseas Chinese Around 1990. Population and Development Review (1994): 631-645.

Rauch, James E., and Vitor Trindade. Ethnic Chinese Networks in International Trade. Review of Economics and Statistics 84, no. 1 (2002): 116-130.

Redding, S.J. and Turner, M.A., Transportation Costs and the Spatial Organization of Economic Activity, in (eds) Gilles Duranton, J. Vernon Henderson and William Strange, Handbook of Urban and Regional Economics, Chapter 20, pages 1339-1398, 2015.

Snyder, L.V., Atan, Z., Peng, P., Rong, Y., Schmitt, A.J. and Sinsoysal, B., 2016. OR/MS Models for Supply Chain Disruptions: A Review. IIE Transactions, 48(2), pp.89-109.

Todo, Y., Nakajima, K. and Matous, P., 2015. How Do Supply Chain Networks Affect the Resilience of Firms to Natural Disasters? Evidence from the Great East Japan Earthquake. Journal of Regional Science, 55(2), pp.209-229.

Topalova, Petia, and Amit Khandelwal. Trade Liberalization and Firm Productivity: the Case of India. Review of Economics and Statistics 93, no. 3 (2011): 995-1009.

Vannoorenberghe, G., Wang, Z. and Yu, Z., 2016. Volatility and Diversification of Exports: Firm-level Theory and Evidence. European Economic Review, 89, pp.216-247.

World Bank. 2016. 2014-2015 West Africa Ebola Crisis: Impact Update. Report.

World Economic Forum, 2012. New Models for Addressing Supply Chain and Transport Risk. Report.

Yeh, C., 2017. Are Firm-Level Idiosyncratic Shocks Important for US Aggregate Volatility?. Working paper, University of Chicago.

Young, A., 2000. The Razor's Edge: Distortions and Incremental Reform in the People's Republic of China. The Quarterly Journal of Economics, 115 (4): 1091-1135.

Young, A., 2005. The Gift of the Dying: the Tragedy of AIDS and the Welfare of Future African Generations. The Quarterly Journal of Economics, 120(2), pp.423-466.

Yu, Miaojie., 2015. Processing Trade, Tariff Reductions and Firm Productivity: Evidence from Chinese Firms. The Economic Journal 125, no. 585: 943-988.

8 Appendix

8.1 Proof of Proposition 1

Suppose there are two firms from region i with different productivities which are denoted as φ_H and φ_L such that $\varphi_H > \varphi_L$. Their sourcing strategies are given by $\mathcal{I}_i(\varphi_H) = \{\{j,k\}, I_{ijk}(\varphi_H) = 1\}$ and $\mathcal{I}_i(\varphi_L) = \{\{j,k\}, I_{ijk}(\varphi_L) = 1\}$ respectively. If $\mathcal{I}_i(\varphi_H) = \mathcal{I}_i(\varphi_L)$, conclusion (a) naturally holds as it implies $\Psi(\varphi_H) = \Psi(\varphi_L)$. On the other hand, if $\mathcal{I}_i(\varphi_H) \neq \mathcal{I}_i(\varphi_L)$, it must be the case that:

$$B_i \varphi_H^{\sigma-1} (\gamma^2 \Psi(\varphi_H))^{\frac{\sigma-1}{\theta}} - \sum_{\{jk\} \in I_{ijk}(\varphi_H)} f_{ijk} > B_i \varphi_H^{\sigma-1} (\gamma^2 \Psi(\varphi_L))^{\frac{\sigma-1}{\theta}} - \sum_{\{jk\} \in I_{ijk}(\varphi_L)} f_{ijk},$$

$$B_i \varphi_L^{\sigma-1} (\gamma^2 \Psi(\varphi_L))^{\frac{\sigma-1}{\theta}} - \sum_{\{jk\} \in I_{ijk}(\varphi_L)} f_{ijk} > B_i \varphi_L^{\sigma-1} (\gamma^2 \Psi(\varphi_H))^{\frac{\sigma-1}{\theta}} - \sum_{\{jk\} \in I_{ijk}(\varphi_H)} f_{ijk}.$$

Combining the two inequalities above, we have

$$B_i \gamma^{\frac{2(\sigma-1)}{\theta}} (\varphi_H^{\sigma-1} - \varphi_L^{\sigma-1}) (\Psi(\varphi_H)^{\frac{\sigma-1}{\theta}} - \Psi(\varphi_L)^{\frac{\sigma-1}{\theta}}) > 0.$$

Since $\varphi_H > \varphi_L$ and $\sigma > 1$, it must be the case that

$$\Psi(\varphi_H) > \Psi(\varphi_L).$$

So conclusion (a) is established.

For conclusion (b), we note that if $\sigma - 1 > \theta$, the profit function specified in problem (3.4) features increasing differences in (I_{ijk}, I_{imn}), with $j \neq m$ or $k \neq n$. It also features increasing differences in (I_{ijk}, φ) for any j and k. Using the Topkis's monotonicity theorem, we conclude that $I_{ijk}(\varphi_H) \geq I_{ijk}(\varphi_L)$ for any $\varphi_H \geq \varphi_L$. It naturally implies $J_i(\varphi_L) \subseteq J_i(\varphi_H)$ and $K_{ij}(\varphi_L) \subseteq K_{ij}(\varphi_H)$.

Then from the definition of $\Theta_j(\varphi) \equiv \sum_{n \in K_{ij}(\varphi)} T_n (\tau_{jn} w_n)^{-\theta}$, given that $K_{ij}(\varphi_L) \subseteq K_{ij}(\varphi_H)$, naturally we have

$$\Theta_j(\varphi_H) \geq \Theta_j(\varphi_L)$$

which is conclusion (c).

8.2 Proof of Proposition 2

Since the inputs are complementary, firms' sourcing strategies follow a pecking order. Suppose the sourcing potential of sourcing routes faced by firms in region i are ranked as $\phi_{i1} \geq \phi_{i2} \geq, ..., \geq \phi_{iN}$. The least productive firm would only source from option 1 and its $HHI_1 = 1$. For two firms with different sourcing strategies such that one is sourcing from n options while the other is sourcing from $n+1$, their HHI for sourcing are $HHI_n = \frac{\sum_{s=1}^{n} \phi_{is}^2}{(\sum_{s=1}^{n} \phi_{is})^2}$ and $HHI_{n+1} = \frac{\sum_{s=1}^{n+1} \phi_{is}^2}{(\sum_{s=1}^{n+1} \phi_{is})^2}$, respectively. Given the pecking order, the firms sourcing from $n+1$ options should be more productive. Moreover, the size of the HHI is given by

$$\begin{aligned}
HHI_{n+1} - HHI_n &= \frac{(\sum_{s=1}^{n+1} \phi_{is}^2)(\sum_{s=1}^{n} \phi_{is})^2 - (\sum_{s=1}^{n} \phi_{is}^2)(\sum_{s=1}^{n+1} \phi_{is})^2}{(\sum_{s=1}^{n} \phi_{is})^2(\sum_{s=1}^{n+1} \phi_{is})^2} \\
&= \frac{\phi_{in+1}^2(\sum_{s=1}^{n} \phi_{is})^2 - \phi_{in+1}^2 \sum_{s=1}^{n} \phi_{is}^2 - 2\phi_{in+1}(\sum_{s=1}^{n} \phi_{is})(\sum_{s=1}^{n} \phi_{is}^2)}{(\sum_{s=1}^{n} \phi_{is})^2(\sum_{s=1}^{n+1} \phi_{is})^2} \\
&= \frac{\phi_{in+1}(\sum_{s=1}^{n} \phi_{is})[\sum_{s=1}^{n} \phi_{in+1}\phi_{is} - \sum_{s=1}^{n} \phi_{is}^2] - \phi_{in+1}^2 \sum_{s=1}^{n} \phi_{is}^2 - \phi_{in+1}(\sum_{s=1}^{n} \phi_{is})(\sum_{s=1}^{n} \phi_{is}^2)}{(\sum_{s=1}^{n} \phi_{is})^2(\sum_{s=1}^{n+1} \phi_{is})^2}.
\end{aligned}$$

Since $\phi_{is} \geq \phi_{in+1}$, it must be the case that $\phi_{is}^2 \geq \phi_{in+1}\phi_{is}, \forall s \leq n$. Then we have $\sum_{s=1}^{n} \phi_{in+1}\phi_{is} \leq \sum_{s=1}^{n} \phi_{is}^2$. Thus the first term in the numerator of the third line in the equation above is non-positive. Given that the other two terms are also negative, the numerator of the third line must be negative. Thus we have

$$HHI_{n+1} < HHI_n,$$

and the concentration of the sourcing strategy tends to lower for more productive firms.

8.3 Proof of Proposition 3

Under input complementarity ($\sigma - 1 > \theta$), the profit function specified in problem (3.4) features increasing difference between (I_{ijk}, ϕ_{imn}) between any $j \neq m$, $k \neq n$. It also features increasing difference between (I_{ijk}, -f_{imn}) between any $j \neq m$, $k \neq n$. Again, using the Topkis's monotonicity theorem, we have $I_{ijk}(\vec{\phi_i}') \geq I_{ijk}(\vec{\phi_i})$ for $\vec{\phi_i}' > \vec{\phi_i}$. Naturally, it implies $J_i(\varphi, \vec{\phi_i}) \subseteq J_i'(\varphi, \vec{\phi_i}')$, $K_{ij}(\varphi, \vec{\phi_i}) \subseteq K_{ij}'(\varphi, \vec{\phi_i}')$. Similarly, we have $I_{ijk}(\vec{f_i}') \geq I_{ijk}(\vec{f_i})$ for $\vec{f_i}' < \vec{f_i}$ which implies $J_i(\varphi, \vec{f_i}) \subseteq J_i'(\varphi, \vec{f_i}')$, $K_{ij}(\varphi, \vec{f_i}) \subseteq K_{ij}'(\varphi, \vec{f_i}')$.

8.4 Proof of Proposition 4

According to Equation (3.3), in case there is any shock to any supplier, the change in unit cost for the firm is given by

$$\widehat{c_i} \equiv \frac{c_i'}{c_i} = \widehat{\Psi}_i(\varphi)^{-\frac{1}{\theta}}.$$

Thus, we have:

$$\frac{\partial ln\widehat{c_i}}{\partial ln\widehat{\Psi}_i(\varphi)} = -\frac{1}{\theta}. \tag{8.9}$$

On the other hand, we have

$$\widehat{\Psi}_i(\varphi) \equiv \frac{\sum_{j\in J'(\varphi),k\in K'(\varphi)} \phi'_{ijk}}{\sum_{j\in J(\varphi),k\in K(\varphi)} \phi_{ijk}}.$$

Suppose $\Omega(\varphi) = J(\varphi) \otimes K(\varphi)$ which is the set of routes picked by the firm before the shock and $\Omega'(\varphi) = J'(\varphi) \otimes K'(\varphi)$ is the one after the shock, and $\mathcal{C}(\varphi) = \Omega \cap \Omega' \neq \oslash$ is the set of routes continued to be used by the firm. The set of new routes used by the firm is denoted as $\mathcal{N}(\varphi) \equiv \Omega' \setminus \mathcal{C}$. Then we have,

$$\begin{aligned}
\widehat{\Psi}_i(\varphi) &= \frac{\sum_{j,k\in\mathcal{C}} \phi'_{ijk} + \sum_{j,k\in\mathcal{N}} \phi'_{ijk}}{\Psi_i(\varphi)} \\
&= \sum_{j,k\in\mathcal{C}} \frac{\phi'_{ijk}}{\phi_{ijk}} \frac{\phi_{ijk}}{\Psi_i(\varphi)} + \sum_{j,k\in\mathcal{N}} \frac{\phi'_{ijk}}{\Psi'_i(\varphi)} \frac{\Psi'_i(\varphi)}{\Psi_i(\varphi)} \qquad (8.10) \\
&= \sum_{j,k\in\mathcal{C}} \widehat{\phi}_{ijk}\chi_{ijk} + \widehat{\Psi}_i(\varphi) \sum_{j,k\in\mathcal{N}} \chi'_{ijk}. \qquad (8.11)
\end{aligned}$$

Rearranging the equation above, we have

$$\widehat{\Psi}_i(\varphi) = \frac{\sum_{j,k\in\mathcal{C}} \chi_{ijk}\widehat{\phi}_{ijk}}{1 - \sum_{j,k\in\mathcal{N}} \chi'_{ijk}}.$$

So one unit change in ϕ_{ijk} translates into $\frac{\chi_{ijk}}{1-\sum_{j,k\in\mathcal{N}} \chi'_{ijk}}$ unit change in $\widehat{\Psi}_i(\varphi)$. Formally, for a small change in x, we know $ln(x) \approx x - 1$. Thus $\widehat{\Psi}_i(\varphi) \approx 1 + \ln(\widehat{\Psi}_i(\varphi))$ and $\widehat{\phi}_{ijk} \approx 1 + \ln(\widehat{\phi}_{ijk})$. Then we have

$$\ln\widehat{\Psi}_i(\varphi) \approx \frac{\sum_{j,k\in\mathcal{C}} \chi_{ijk} \ln\widehat{\phi}_{ijk}}{1 - \sum_{j,k\in\mathcal{N}} \chi'_{ijk}} + \frac{\sum_{j,k\in\mathcal{C}}(\chi_{ijk} - \chi'_{ijk})}{1 - \sum_{j,k\in\mathcal{N}} \chi'_{ijk}}$$

which implies

$$\frac{\partial \ln\widehat{\Psi}_i(\varphi)}{\partial \ln\widehat{\phi}_{ijk}} \approx \frac{\chi_{ijk}}{1 - \sum_{j,k\in\mathcal{N}} \chi'_{ijk}}.$$

Finally, since $\phi_{ijk} = A_j T_k(\tau_{ijk} w_j w_k)^{-\theta}$, we have $\frac{\partial \ln \widehat{\phi}_{ijk}}{\partial \ln \widehat{\tau}_{ijk}} = -\theta$. This implies that the pass-through of cost shock $\widehat{\tau}_{ijk}$ to marginal cost of the firm is given by:

$$\begin{aligned}\frac{\partial ln\widehat{c}_i}{\partial \ln \widehat{\tau}_{ijk}} &= \frac{\partial ln\widehat{c}_i}{\partial ln\widehat{\Psi}_i(\varphi)} \frac{\partial \ln \widehat{\Psi}_i(\varphi)}{\partial \ln \widehat{\phi}_{ijk}} \frac{\partial \ln \widehat{\phi}_{ijk}}{\partial \ln \widehat{\tau}_{ijk}} \\ &\approx \frac{\chi_{ijk}(\varphi)}{1 - \sum_{j,k\in\mathcal{N}} \chi'_{ijk}(\varphi)}.\end{aligned}$$

The proof of conclusion (b) has two steps. First, from proposition 1, we know that sourcing capabilities $\Psi(\varphi)$ is an increasing function of productivity φ. Thus the probability of sourcing from route imn $\chi_{ijk}(\varphi) = \frac{A_j T_k(\tau_{ijk} w_j w_k)^{-\theta}}{\Psi(\varphi)}$ is decreasing with φ. Second, for the denominator $1 - \sum_{j,k\in\mathcal{N}} \chi_{ijk}$, according to Proposition 3, we have $\sum_{j,k\in\mathcal{N}} \chi_{ijk} = 0$ in the case of adverse shocks, and $\sum_{j,k\in\mathcal{N}} \chi_{ijk} \geq 0$ in the case of favourable shock. Alternatively, this result can be derived by studying how the productivity cut-offs respond to the shocks as follows.

In the case of an adverse shock such that τ_{ijk} increases, it can be shown that no firms would like to increase the number of sourcing routes. To show that, we know that there is a pecking order in the case of complementarity and we could rank different sourcing routes according to their sourcing potential. The most appealing option would be sourced by all firms. This is option 1. The least appealing option would only be sourced by the most productive firms. This is option N. Suppose the productivity cut-offs for these different options are $\tilde{\varphi}_{i1} \leq \tilde{\varphi}_{i2} \leq, ..., \leq \tilde{\varphi}_{iN}$ and suppose the route which is shocked currently ranked at r. We can know that the cut-offs are determined by

$$\tilde{\varphi}_{i1}^{\sigma-1} = \frac{w_i f_{i1}}{\gamma^{(\frac{2(\sigma-1)}{\theta}} B_i \phi_{i1}^{\frac{(\sigma-1)}{\theta}}}$$

$$\tilde{\varphi}_{in}^{\sigma-1} = \frac{w_i f_{in}}{\gamma^{(\frac{2(\sigma-1)}{\theta}} B_i((\sum_{l=1}^{n} \phi_{il})^{\frac{(\sigma-1)}{\theta}} - (\sum_{l=1}^{n-1} \phi_{il})^{\frac{(\sigma-1)}{\theta}})}, n > 1.$$

when the trade costs using route r increases, it will not affect cutoffs $\tilde{\varphi}_{i1}, \tilde{\varphi}_{i2}, ..., \tilde{\varphi}_{ir-1}$. Firms with productivity lower than $\tilde{\varphi}_{ir-1}$ will keep their sourcing routes as they are. However, as τ_{ir} increases, the sourcing potential of route r: ϕ_{ir} will decrease. This will decrease the difference between sourcing capabilities through n routes v.s. $n-1$ routes for $n \geq r$ as $\frac{\sigma-1}{\theta} \geq 1$. [59] Then for all $n \geq r$, we have $\tilde{\varphi}_{in}$ increases. This is illustrated in Figure 8 (a). Thus no firms would like to add sourcing routes. Instead, they would decrease the number of sourcing routes. So we

[59] It can be shown that $f(x) = f(x+a)^{\frac{\sigma-1}{\theta}} - x^{\frac{\sigma-1}{\theta}}$ is an increasing function of x as $\frac{\sigma-1}{\theta} \geq 1$.

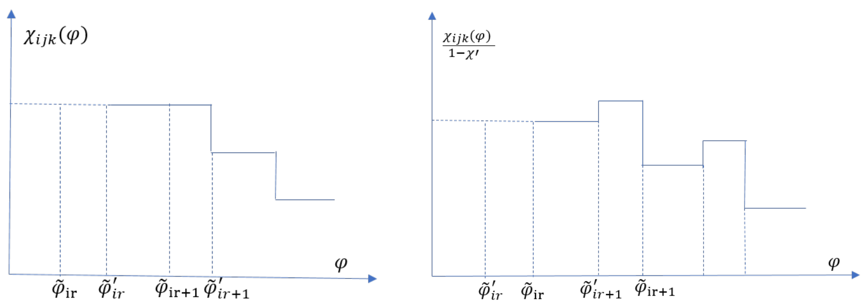
have $1-\sum_{j,k\in\mathcal{N}}\chi_{ijk}=1$ for all firms and

$$\frac{\partial ln\widehat{c_i}}{\partial\ln\widehat{\tau}_{ijk}} \quad\approx\quad \chi_{ijk}(\varphi).$$

which declines with φ. So we have

$$\frac{\partial^2\ln(\widehat{c_i(\varphi)})}{\partial\ln(\widehat{\tau_{ijk}})\partial\varphi}=\frac{\partial\chi_{ijk}(\varphi)}{\partial\varphi}\leq 0.$$

$$\frac{\partial^2\ln(\widehat{c_i(\varphi)})}{\partial\ln(\widehat{\tau_{ijk}})\partial\phi_{ijk}}=\frac{\partial\chi_{ijk}(\varphi)}{\partial\phi_{ijk}}=\frac{1}{\Psi_i(\varphi)}>0.$$



(a) Adverse Shocks (b) Favourable Shocks

Figure 8: Pass-through of shocks

In the case that τ_{ijk} decreases. Following the previous case, it is easy to see that $\tilde{\varphi}_{i1},\tilde{\varphi}_{i2},...,\tilde{\varphi}_{ir-1}$ are not affected and firms with productivity $\varphi\leq\tilde{\varphi}_{ir-1}$ do not change their sourcing strategies. Intuitively, they do not include r in their sourcing options and are not affected by the cost shock. On the other hand, $\tilde{\varphi}_{in}$ will decrease to $\tilde{\varphi}'_{in}$ for all $n\geq r$. Then for firms with productivity within $[\tilde{\varphi}'_{i,n+1},\tilde{\varphi}_{i,n+1}]$, they would like to include $n+1$ in their sourcing strategies. The pass-through of the shock would be

$$\frac{\partial ln\widehat{c_i}}{\partial\ln\widehat{\tau}_{ijk}} \quad\approx\quad \frac{\chi_{ir}(\varphi)}{1-\chi_{i,n+1}(\varphi)'}.$$

Firms with productivity in $[\tilde{\varphi}_{in},\tilde{\varphi}'_{i,n}]$ fix their sourcing strategies and we still have

$$\frac{\partial ln\widehat{c_i}}{\partial\ln\widehat{\tau}_{ijk}} \quad\approx\quad \chi_{ijk}(\varphi).$$

In this case, the pass-through is not universally declining with productivity as illustrated by

Figure 8 (b).

8.5 Proof of Proposition 5

The gravity equation at firm level determining the trade flow is given by Equation (3.5). Facing a supply shock, the change in trade flow is determined by

$$\begin{aligned}\widehat{M}_{ijk}(\varphi) &\equiv \frac{M'_{ijk}(\varphi)}{M_{ijk}(\varphi)} = \widehat{\Psi}_i(\varphi)^{\frac{(\sigma-1)}{\theta}}\widehat{\chi}_{ijk}(\varphi) \\ &= \widehat{\Psi}_i(\varphi)^{\frac{(\sigma-1)}{\theta}-1}\widehat{\phi}_{ijk},\end{aligned}$$

which implies $\ln \widehat{M}_{ijk}(\varphi) = (\frac{\sigma-1}{\theta} - 1)\ln \widehat{\Psi}_i(\varphi) + \ln \widehat{\phi}_{ijk}$. From the previous proof, we know that for an adverse shock

$$\frac{\partial \ln \widehat{\Psi}_i(\varphi)}{\partial \ln \widehat{\phi}_{ijk}} \approx \chi_{ijk}.$$

And since $\frac{\partial \ln \widehat{\phi}_{i.mn}}{\partial \ln \widehat{\tau}_{i.mn}} = -\theta$, we have

$$-\frac{\partial \ln \widehat{M}_{ijk}(\varphi)}{\partial \ln \widehat{\tau}_{imn}} = \begin{cases} \theta + (\sigma - 1 - \theta)\chi_{imn}, \text{ if m=j, n=k} \\ (\sigma - 1 - \theta)\chi_{imn}, \text{otherwise.}\end{cases}$$

This establishes conclusion (a). Under input complementarity, as I mentioned in the previous proof, the sourcing probability $\chi_{ijk}(\varphi)$ of each sourcing route is weakly decreasing in firm productivity φ. Then, the size of the pass-through should follow the same pattern under input complementarity. This establishes conclusion (b).

8.6 Proof of Proposition 6

From the proof of Proposition 4, we know that the change in sourcing capability $\Psi = \sum \phi_r$ for a particular firm is given by [60]

$$\widehat{\Psi}(\varphi) = \frac{\sum\limits_{r\in\mathcal{C}(\varphi)} \chi_r(\varphi)\widehat{\phi_r}}{1 - \sum\limits_{r\in\mathcal{N}} \chi'_r(\varphi)},$$

where $\mathcal{C}(\varphi) \subset \Omega(\varphi)$ and $\mathcal{N} \subset \Omega^{'}(\varphi)$ are the sets of continued and new sourcing route for the firm respectively while $\Omega(\varphi)$ and $\Omega^{'}(\varphi)$ are the set of sourcing routes before and after the shocks.

[60] To simplify the notation, I omit the location subscript i.

They all depend firm level productivity φ. I further simplify the notations as

$$\widehat{\Psi}(\varphi)=\sum_{r\in\Omega(\varphi)}\chi_r(\varphi)\Delta_r,$$

while $\Delta_r=\widehat{\phi_r}\delta_r(\varphi,\widehat{\phi})$ with $\delta_r(\varphi,\widehat{\phi})$ being an indicator function defined as

$$\delta_r(\varphi,\widehat{\phi})=\begin{cases}\frac{1}{1-\sum\limits_{r\in\mathcal{N}}\chi_r'(\varphi)}, if\ r\in\mathcal{C}(\varphi,\widehat{\phi});\\ 0,\ \ otherwise,\end{cases}$$

which captures to extensive margin shock of sourcing capabilities. Under the assumption that Δ_r has the same variance ξ^2 across sourcing routes, we have

$$\begin{aligned}var(\widehat{\Psi}(\varphi)) &= var(\sum_{r\in\Omega(\varphi)}\chi_r(\varphi)\Delta_r)\\ &= \sum_{r\in\Omega(\varphi)}\chi_r(\varphi)^2 var(\Delta_r)+\sum_{m\neq n,m,n\in\Omega(\varphi)}\chi_m(\varphi)\chi_n(\varphi)cov(\Delta_m,\Delta_n)\\ &= \xi^2(\sum_{r\in\Omega(\varphi)}\chi_r(\varphi)^2+\sum_{m\neq n,m,n\in\Omega(\varphi)}\chi_m(\varphi)\chi_n(\varphi)\rho_{mn})\\ &\leq \xi^2.\end{aligned}$$

The last inequality holds because $(\sum\limits_{r\in\Omega(\varphi)}\chi_r(\varphi))^2=\sum\limits_{r\in\Omega(\varphi)}\chi_r(\varphi)^2+\sum\limits_{m\neq n,m,n\in\Omega(\varphi)}\chi_m(\varphi)\chi_n(\varphi)=1$. As long as the correlation of shocks $\rho_{ij}\equiv\frac{cov(\Delta_m,\Delta_n)}{\xi^2}<1$ for any i and j, that is the shocks are not perfectly correlated across sourcing routes, we have $var(\widehat{\Psi}(\varphi))<\xi^2$. On the other hand, if firms are under sourcing autarky, firms are subject to local shocks with volatility ξ^2. This establishes conclusion (a).

If the shocks are i.i.d. such that $\rho_{mn}=0$, we have:

$$\begin{aligned}var(\widehat{\Psi}(\varphi)) &= var(\sum_{r\in\Omega(\varphi)}\chi_r(\varphi)\Delta_r)\\ &= \xi^2\sum_{r\in\Omega(\varphi)}\chi_r(\varphi)^2\\ &= \xi^2 HHI(\varphi).\end{aligned}$$

From Proposition 2, we know that HHI decrease weakly with firm productivity. Then the volatility of firms' sourcing capabilities should decrease weakly with firm productivity. For firm revenue given by $R_i(\varphi) = B_i \varphi^{\sigma-1} \gamma^{\frac{2(\sigma-1)}{\theta}} \Psi_i(\varphi)^{\frac{\sigma-1}{\theta}}$, thus we have

$$\widehat{R}_i(\varphi) = \widehat{\Psi}_i(\varphi)^{\frac{\sigma-1}{\theta}}.$$

B_i does not show up as we only care about cost shock through inputs. Using the delta method, we have

$$\begin{aligned} var(\widehat{R}_i(\varphi)) &\approx [\frac{\partial \widehat{R}_i(\widehat{\Psi}_i(\varphi))}{\partial \widehat{\Psi}_i(\varphi)}\Big|_{\widehat{\Psi}_i(\varphi)=E[\widehat{\Psi}_i(\varphi)]}]^2 var(\widehat{\Psi}_i(\varphi)) \\ &= \frac{(\sigma-1)^2}{\theta^2} E[\widehat{\Psi}_i(\varphi)]^{\frac{2(\sigma-1-\theta)}{\theta}} var(\widehat{\Psi}_i(\varphi)). \end{aligned}$$

Under the assumption that Δ_r is i.i.d., the expectation for $\widehat{\Psi}_i(\varphi) : E[\widehat{\Psi}_i(\varphi)] = E[\sum_{r \in \Omega(\varphi)} \chi_r(\varphi) \Delta_r] = \sum_{r \in \Omega(\varphi)} \chi_r(\varphi) E[\Delta_r] = E[\Delta_r]$ is a constant. The last equality holds as $\sum_{r \in \Omega(\varphi)} \chi_r(\varphi) = 1$. Since $var(\widehat{\Psi}(\varphi)) = \xi^2 HHI(\varphi)$, we have

$$var(\widehat{R}_i(\varphi)) \propto \xi^2 HHI(\varphi),$$

which declines weakly with firm productivity φ in the same way as $HHI(\varphi)$. This establishes conclusion (b).

Under universal importing, the change to the sourcing capabilities of a firm is given by

$$\widehat{\Psi} = \sum_{r \in \Omega} \chi_r \widehat{\phi_r},$$

where Ω is set of available sourcing routes to all importers. Given the universal importing assumption, Ω is the same for each firm, so is the sourcing intensity χ_r. Then the variance of $\widehat{\Psi}$ should be the same for all importers. So is firm revenue. This establishes conclusion (c).

8.7 Proof of Proposition 7

From Equation (3.3), the change in marginal costs in response to sourcing potentials is given by

$$\widehat{c}_i(\varphi) = \widehat{\Psi}_i(\varphi)^{-\frac{1}{\theta}} \tag{8.12}$$

which is inversely related to the change in the sourcing capability of the firm. On the other hand, from the proof of Proposition 4, the change in sourcing capability to an adverse shock is related to change sourcing potential and pre-shock sourcing probability as

$$\widehat{\Psi}_i(\varphi) = \sum_{j,k\in\mathcal{C}} \chi_{ijk}\widehat{\phi}_{ijk}. \tag{8.13}$$

under input complementarity. Although $\widehat{\phi}_{ijk}$ is still not observable, according to Equation (3.5), the change in the trade flow is given by

$$\begin{aligned}\widehat{M}_{ijk}(\varphi) &= \widehat{\Psi}_i(\varphi)^{\frac{\sigma-1}{\theta}}\frac{\widehat{\phi}_{ijk}(\phi)}{\widehat{\Psi}_i(\varphi)} \\ &= \widehat{\Psi}_i(\varphi)^{\frac{\sigma-1}{\theta}-1}\widehat{\phi}_{ijk}(\phi)\end{aligned}$$

which implies

$$\widehat{\phi}_{ijk}(\phi) = \frac{\widehat{M}_{ijk}(\varphi)}{\widehat{\Psi}_i(\varphi)^{\frac{\sigma-1}{\theta}-1}}.$$

Substitute the equation above into Equation (8.13), we have

$$\widehat{\Psi}_i(\varphi) = (\sum_{j\times k\in C}\chi_{ijk}(\varphi)\widehat{M}_{ijk}(\varphi))^{\frac{\theta}{\sigma-1}},$$

together with Equation (8.12), immediately we know

$$\widehat{c}_i(\varphi) = (\sum_{j\times k\in C}\chi_{ijk}(\varphi)\widehat{M}_{ijk}(\varphi))^{\frac{1}{1-\sigma}}.$$

8.8 Tradable Final Goods and Demand Shocks

Suppose final goods are tradable. Exporting to market k through customs district j incurs a variable iceberg trade cost τ^X_{ijk}, and a fixed cost in terms of f^X_{ijk} unit of labor from region i. Then firms' profit function in Equation (3.4) now becomes:

$$\max_{I_{ijk},\, I^X_{ijk}\in\{0,1\}} \pi(\varphi, \{I_{ijk}\}, \{I^X_{ijk}\}) = \varphi^{\sigma-1}(\gamma^2\Psi_i(\varphi))^{\frac{\sigma-1}{\theta}}B_i(\varphi) - w_i\sum_{j=1,k=1}^{J,K} I_{ijk}f_{ijk} - w_i\sum_{j=1,k=1}^{J,K} I^X_{ijk}f^X_{ijk}$$

where I_{ijk} and I^X_{ijk} are indicator variables for import and export through route jk respectively, and $B_i(\varphi) \equiv \sum_{j=1,k=1}^{J,K} I^X_{ijk}(\tau^X_{ijk})^{1-\sigma}B_k$ is the demand shifter. The model features increasing difference in (I^X_{ijk}, φ), so more productive firms tend to export to more places. It also features increasing difference in $(I^X_{ijk}, \Psi_i(\varphi))$, thus any reduction in trade costs would lead firms to expand their export along the extensive margin, *vice versa* if trade costs increase.

The gravity equation at firm level determining the import flow is still given by Equation (3.5) except that the demand shifter B_i is now firm specific. Suppose $\widehat{\tau}^X_{ijk} = \widehat{\tau}_{ijk}$, thus the cost shock affects imports and exports along the same route at the same time, then

$$\widehat{M}_{ijk}(\varphi) = \widehat{\Psi}_i(\varphi)^{\frac{(\sigma-1)}{\theta}-1}\widehat{\phi}_{ijk}\widehat{B}_i(\varphi).$$

Compared with the proof in Appendix 8.5, there is an extra term $\widehat{B}_i(\varphi)$ which captures the demand shock given by:

$$\begin{aligned}
\widehat{B}_i(\varphi) &= \frac{\sum_{j,k\in\mathcal{C}^{\mathcal{X}}(\varphi)} b'_{ijk} + \sum_{j,k\in\mathcal{N}^{\mathcal{X}}(\varphi)} b'_{ijk}}{B_i(\varphi)} \\
&= \sum_{j,k\in\mathcal{C}^{\mathcal{X}}(\varphi)} \frac{b'_{ijk}}{b_{ijk}}\frac{b_{ijk}}{B_i(\varphi)} + \sum_{j,k\in\mathcal{N}^{\mathcal{X}}(\varphi)} \frac{b'_{ijk}}{B'_i(\varphi)}\frac{B'_i(\varphi)}{B_i(\varphi)} \\
&= \sum_{j,k\in\mathcal{C}^{\mathcal{X}}(\varphi)} \widehat{b}_{ijk}\mu_{ijk} + \widehat{B}_i(\varphi) \sum_{j,k\in\mathcal{N}^{\mathcal{X}}(\varphi)} \mu'_{ijk},
\end{aligned}$$

where $\mathcal{C}^{\mathcal{X}}(\varphi)$ is the set of destinations that the firm continues to serve after the shock, and $\mathcal{N}^{\mathcal{X}}(\varphi)$ is the set of destinations that are newly included. Rearranging the equation above, we have

$$\widehat{B}_i(\varphi) = \frac{\sum_{j,k\in\mathcal{C}^{\mathcal{X}}} \mu_{ijk}\widehat{b}_{ijk}}{1 - \sum_{j,k\in\mathcal{N}^{\mathcal{X}}} \mu'_{ijk}},$$

where $\mu_{ijk}(\varphi) \equiv \frac{b_{ijk}}{B_i(\varphi)}$ is the intensity of exporting through route jk, and $b_{ijk} \equiv (\tau^X_{ijk})^{1-\sigma} B_k$ is the residual demand for route jk. For negative shocks on trade costs, as argued above, firms would like to reduce exports along the extensive margin. Thus $\sum_{j,k\in\mathcal{N}^{\mathcal{X}}(\varphi)} \mu'_{ijk} = 0$ and

$$\widehat{B}_i(\varphi) = \sum_{j,k\in\mathcal{C}^{\mathcal{X}}(\varphi)} \mu_{ijk}\widehat{b}_{ijk}.$$

Then we have

$$\begin{aligned}
\frac{\partial \ln \widehat{B_i}}{\partial \ln \widehat{\tau}_{ijk}} &= \frac{\partial \ln \widehat{B_i}}{\partial \ln \widehat{b}_{ijk}}\frac{\partial \ln \widehat{b}_{ijk}}{\partial \ln \widehat{\tau}_{ijk}} \\
&= (1-\sigma)\mu_{ijk}(\varphi),
\end{aligned}$$

combined with the fact that $\frac{\partial \ln \widehat{M}_{ijk}(\varphi)}{\partial \ln \widehat{\tau}_{imn}} = (\frac{\sigma-1}{\theta} - 1)\frac{\partial \ln \widehat{\Psi}_i(\varphi)}{\partial \ln \widehat{\tau}_{imn}} + \frac{\partial \ln \widehat{\phi}_{ijk}}{\partial \ln \widehat{\tau}_{imn}} + \frac{\partial \ln \widehat{B}_i(\varphi)}{\partial \ln \widehat{\tau}_{imn}}$, we have

$$-\frac{\partial \ln \widehat{M}_{ijk}(\varphi)}{\partial \ln \widehat{\tau}_{imn}} = \begin{cases} \theta + (\sigma - 1 - \theta)\chi_{imn}(\varphi) + (\sigma - 1)\mu_{imn}(\varphi), \text{ if m=j, n=k,} \\ (\sigma - 1 - \theta)\chi_{imn}(\varphi) + (\sigma - 1)\mu_{imn}(\varphi), \text{otherwise.} \end{cases}$$

9 Additional Tables and Figures

9.1 Additional Tables

9.1.1 Output Volatility and sourcing diversification

This subsection provides robustness tests on the firms' output volatility and input diversification. The first exercise looks at the volatility of firms' export which is generated out of a relatively long time series of firms' quarterly export. I then regress the volatility of exports on the number of sourcing routes, controlling firm age, firm size and output diversification captured by the number of exporting routes. The results are shown in Table A2. As can be seen from the table, firms with more diversified sourcing strategies continue to have lower volatility in exports.

Table A1: Volatility of sales

Dependent variable: ln(sales volatility)	(1)	(2)	(3)	(4)	(5)
Sourcing Diversification measured by HHI	0.288***	0.369***	0.270***	0.218***	0.180***
	(0.0315)	(0.0306)	(0.0310)	(0.0328)	(0.0414)
ln age of firm		-0.207***	-0.176***	-0.188***	-0.228***
		(0.00750)	(0.00763)	(0.00789)	(0.0154)
ln employment			-0.0708***	-0.0448***	-0.0288***
			(0.00472)	(0.00592)	(0.00928)
ln TFP				-0.0438***	-0.0360***
				(0.00665)	(0.0105)
ln number of exporting routes				-0.00991**	-0.0159**
				(0.00462)	(0.00733)
ln number of imported products					-0.0158**
					(0.00697)
Industry FE	Y	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y	Y
R^2	0.0556	0.0806	0.0878	0.0895	0.0958
No. of observations	44454	44451	44451	44432	14356

Notes: Sales volatility is the variance of growth rate during 1999-2007. Herfindahl-Hirschman Index (HHI) is constructed over the shares of inputs sourced from different trade routes and averaged across years. For non-importers, it is assigned as 1. For exporting routes, it is assigned as 1 for non-exporters. Column (5) only includes importers. The numbers in the parentheses are standard error clustered at industry and prefecture level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table A2: Volatility of exports

Dependent variable: ln(exports volatility)	(1)	(2)	(3)	(4)	(5)
Sourcing Diversification measured by HHI	0.732***	0.740***	0.663***	0.685***	0.519***
	(0.0669)	(0.0665)	(0.0681)	(0.0670)	(0.0844)
ln age of firm		-0.0703*	-0.000882	0.00268	-0.0119
		(0.0385)	(0.0386)	(0.0386)	(0.0410)
ln employment			-0.0950***	-0.101***	-0.0881***
			(0.0153)	(0.0209)	(0.0218)
ln TFP				0.0462**	0.0650***
				(0.0215)	(0.0229)
ln number of exporting routes				-0.0514***	-0.0552***
				(0.0166)	(0.0169)
ln number imported products					-0.0586***
					(0.0177)
Industry FE	Y	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y	Y
R^2	0.0846	0.0853	0.0931	0.0957	0.0987
No. of observations	5887	5887	5887	5884	5716

Notes: Volatility of exports is the variance of quarterly exports growth rate between 2000 and 2006. Herfindahl-Hirschman Index (HHI) is constructed over the shares of inputs sourced from different trade routes and averaged across years. Only firms that are both importer and exporter are included in column (5). The numbers in the parentheses are standard error clustered at industry and prefecture level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

9.1.2 Multi-customs-district importer premium

This section presents results on the various robustness checks on the multi customs district premium. First, I show that this is not a phenomenon particular to year 2006. Table A4 presents results year 2000. The premium is quite similar. Additional checks are shown in Table A5. First, an alternative measure of multi-plant firm is used. The measure is a variable in the CAIS data called "Dan1wei4shu4liang4" in Pinyin which means number of production unit. This is not the number of plants that a firm has but multiple-plant firms should have more production units. The results is shown in column (1) and (2). There is worry that some regions might have place-based policy such as processing trade zone. It might induce firms importing from certain places. I exclude firms that purely engage in processing imports.[61] The result is shown in column (3) and (4). Finally, given that Guangdong Province is divided into 7 custom areas, significantly more than other provinces. To address the concern that the result is driven importers from Guangdong, I exclude importers from Guangdong. The result is presented in column (5) and (6). The multi-customs-district premium remains robust.

[61]Processing import is defined as pure assembly (14 in the 2-digit shipment id code) and processing with imported materials (15 in the 2-digit shipment id code).

Table A3: Multi Customs District Premium 2006

	(1) Sales	(2) Sales	(3) Sales	(4) Sales	(5) Import	(6) labor productivity	(7) TFP
2 customs districts	0.593***	0.151***	0.118***	0.107***	0.599***	0.109***	0.0906***
	(0.0187)	(0.0180)	(0.0134)	(0.0133)	(0.0304)	(0.0146)	(0.0151)
3 customs districts	1.115***	0.381***	0.263***	0.229***	0.828***	0.246***	0.269***
	(0.0378)	(0.0353)	(0.0250)	(0.0247)	(0.0496)	(0.0285)	(0.0297)
4 customs districts	1.662***	0.627***	0.459***	0.389***	0.999***	0.458***	0.442***
	(0.0763)	(0.0704)	(0.0489)	(0.0491)	(0.0825)	(0.0593)	(0.0613)
5+ customs districts	2.226***	0.989***	0.710***	0.558***	1.517***	0.589***	0.585***
	(0.118)	(0.105)	(0.0768)	(0.0755)	(0.131)	(0.101)	(0.0961)
ln # of import countries		0.691***	0.280***	0.272***	1.620***	0.140***	0.424***
		(0.0117)	(0.00857)	(0.00861)	(0.0198)	(0.00843)	(0.00926)
ln Employment			0.775***	0.769***	0.359***		
			(0.00591)	(0.00595)	(0.0117)		
Industry FE	Y	Y	Y	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y	Y	Y	Y
Ownership FE	Y	Y	Y	Y	Y	Y	Y
Multi-plant FE	N	N	N	Y	Y	Y	Y
R^2	0.291	0.419	0.681	0.683	0.569	0.410	0.403
No. of observations	37589	37589	37589	37571	37572	36324	36211

Notes: The estimation method is OLS with high dimensional FE using the Stata command *reghdfe* written by Correia (2015). It eliminates singletons nested within fixed effects. From column (3) on, we control the multi-plant firms using the measure of firm unit number in the data. Industry fixed effects is control at the 4-digit CIC level. Ownership fixed effects is controlled using the registered firm type which distinguishes firms between state owned enterprises, private owned enterprises and foreign invested enterprises. The numbers in the parentheses are standard error clustered at industry and prefecture level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table A4: Multi customs district premium 2000

	(1) Sales	(2) Sales	(3) Sales	(4) Sales	(5) Import	(6) labor productivity	(7) TFP
2 customs districts	0.622***	0.221***	0.143***	0.135***	0.738***	0.103***	0.128***
	(0.0267)	(0.0253)	(0.0190)	(0.0190)	(0.0442)	(0.0234)	(0.0241)
3 customs districts	1.058***	0.377***	0.325***	0.297***	1.023***	0.293***	0.261***
	(0.0532)	(0.0491)	(0.0380)	(0.0379)	(0.0784)	(0.0501)	(0.0466)
4 customs districts	1.587***	0.660***	0.429***	0.424***	1.268***	0.402***	0.502***
	(0.132)	(0.122)	(0.0986)	(0.0950)	(0.137)	(0.0980)	(0.101)
5+ customs districts	2.179***	1.050***	0.682***	0.463***	1.087***	0.619***	0.927***
	(0.236)	(0.227)	(0.158)	(0.149)	(0.310)	(0.196)	(0.207)
ln # of import countries		0.611***	0.285***	0.278***	1.489***	0.144***	0.386***
		(0.0129)	(0.00992)	(0.00987)	(0.0274)	(0.0113)	(0.0118)
ln Employment			0.708***	0.707***	0.343***		
			(0.00856)	(0.00855)	(0.0181)		
Industry FE	Y	Y	Y	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y	Y	Y	Y
Ownership FE	Y	Y	Y	Y	Y	Y	Y
Multi-plant FE	N	N	N	Y	Y	Y	Y
R^2	0.355	0.458	0.676	0.677	0.526	0.331	0.317
No. of observations	17984	17984	17984	17974	17980	17444	17407

Notes: The estimation method is OLS with high dimensional FE using the Stata command *reghdfe* written by Correia (2015). It eliminates singletons nested within fixed effects. From column (3) on, we control the multi-plant firms using the measure of firm unit number in the data. Industry fixed effects is control at the 4-digit CIC level. Ownership fixed effects is control using the registered firm type which distinguishes firms between state owned enterprises, private owned enterprises and foreign invested enterprises. The numbers in the parentheses are standard error clustered at industry and prefecture level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table A5: Robustness of Multi-customs-district Premium 2006

	(1) Sales	(2) TFP	(3) Sales	(4) TFP	(5) Sales	(6) TFP
2 customs districts	0.116***	0.104***	0.0870***	0.0646***	0.0401***	0.0397**
	(0.0134)	(0.0150)	(0.0149)	(0.0174)	(0.0155)	(0.0182)
3 customs districts	0.256***	0.304***	0.190***	0.205***	0.123***	0.174***
	(0.0249)	(0.0299)	(0.0261)	(0.0326)	(0.0279)	(0.0345)
4 customs districts	0.453***	0.517***	0.331***	0.360***	0.223***	0.277***
	(0.0490)	(0.0607)	(0.0499)	(0.0628)	(0.0532)	(0.0677)
5+ customs districts	0.694***	0.740***	0.458***	0.414***	0.421***	0.463***
	(0.0772)	(0.0996)	(0.0766)	(0.0980)	(0.0781)	(0.101)
ln # of import countries	0.280***	0.435***	0.314***	0.465***	0.321***	0.450***
	(0.00858)	(0.00924)	(0.00967)	(0.0103)	(0.00958)	(0.0104)
ln Employment	0.771***		0.776***		0.764***	
	(0.00598)		(0.00661)		(0.00716)	
Industry FE	Y	Y	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y	Y	Y
Ownership FE	Y	Y	Y	Y	Y	Y
Multi-plant FE	Y	Y	Y	Y	Y	Y
R^2	0.681	0.401	0.702	0.417	0.706	0.406
No. of observations	37571	36213	26265	25168	23896	22940

Notes: Column (1) and (2) use alternative measure of mulit-plant firm. Column (3) and (4) exclude pure processing importers. Column (5) and (6) exclude importers from Guangdong province. The numbers in the parentheses are standard error clustered at industry and prefecture level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table A6: Area with local transmission of SARS

Country	Area	From	To
China	Beijing	02-Mar-03	18-Jun-03
China	Guangdong	16-Nov-02	07-Jun-03
China	Hebei	19-Apr-03	10-Jun-03
China	Hubei	17-Apr-03	26-May-03
China	Inner Mongolia	04-Mar-03	03-Jun-03
China	Jilin	01-Apr-03	29-May-03
China	Jiangsu	19-Apr-03	21-May-03
China	Shanxi	08-Mar-03	13-Jun-03
China	Shaanxi	12-Apr-03	29-May-03
China	Tianjin	16-Apr-03	28-May-03
China	Hong Kong	15-Feb-03	22-Jun-03
China	Taiwan	25-Feb-03	05-Jul-03
Canada	Greater Toronto Area	23-Feb-03	02-Jul-03
Canada	New Westminster	28-Mar-03	05-May-03
Mongolia	Ulaanbaatar	05-Apr-03	09-May-03
Philippines	Manila	06-Apr-03	19-May-03
Singapore	Singapore	25-Feb-03	31-May-03
Vietnam	Hanoi	23-Feb-03	27-Apr-03

Notes: This table lists the areas identified by the WHO as regions with local transmission of SARS.

Table A7: Resilience of processing importers: partial processing traders

Dependent Variable:	Processing with inputs		Pure Assembly	
firm import by route $ln(imp_{ijk})$	(1)	(2)	(3)	(4)
pre shock sourcing intensity	9.456***	9.510***	7.542***	7.532***
	(0.117)	(0.114)	(0.222)	(0.220)
trade route hit by SARS=1	-0.0650**	-0.0248	-0.0596	-0.0673
	(0.0281)	(0.0296)	(0.0757)	(0.0816)
trade route hit by SARS=1 x pre shock sourcing intensity		-0.662***		0.150
		(0.138)		(0.452)
other routes hit by SARS=1		-0.0103		0.0204
		(0.0249)		(0.0705)
firm-time FE	Y	Y	Y	Y
industry FE	Y	Y	Y	Y
ownership type FE	Y	Y	Y	Y
origin FE	Y	Y	Y	Y
destination FE	Y	Y	Y	Y
customs area FE	Y	Y	Y	Y
R^2	0.470	0.470	0.487	0.487
No. of observations	1385724	1385724	78896	78896

Notes: Partial processing traders are importers which partially participate in processing trade. Column (1) and (2) use the sample of importers that engage both in Processing with Inputs (PI) and ordinary imports but not Pure Assembly (PA). Column (3) and (4) use the sample of importers that engage both in PA and ordinary imports but not PI. The numbers in the parentheses are standard errors clustered at region-industry level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table A8: Resilience of importers: exporters and non-exporters

Dependent Variable:	Importers which do not export		Importers which export	
firm import by route $ln(imp_{ijk})$	(1)	(2)	(3)	(4)
pre shock sourcing intensity	7.863***	7.985***	9.547***	9.574***
	(0.248)	(0.244)	(0.0987)	(0.0976)
trade route hit by SARS=1	0.00956	0.0733	-0.0819***	-0.0626**
	(0.0623)	(0.0646)	(0.0238)	(0.0249)
trade route hit by SARS=1 x pre shock sourcing intensity		-1.665***		-0.375***
		(0.361)		(0.128)
other routes hit by SARS=1		-0.0407		0.00474
		(0.0589)		(0.0205)
firm-time FE	Y	Y	Y	Y
industry FE	Y	Y	Y	Y
ownership type FE	Y	Y	Y	Y
origin FE	Y	Y	Y	Y
destination FE	Y	Y	Y	Y
customs area FE	Y	Y	Y	Y
R^2	0.549	0.550	0.467	0.467
No. of observations	153627	153627	1866083	1866083

Notes: This table examines the resilience of importers which export and those that do not. Column (1) and (2) include importers that do not export. Column (3) and (4) include importers that are also exporters. The numbers in the parentheses are standard errors clustered at region-industry level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table A9: Resilience of importers: single-location and multi-location importers

Dependent Variable:	Single-location firm		Multi-location firm	
firm import by route $ln(imp_{ijk})$	(1)	(2)	(3)	(4)
pre SARS sourcing intensity	9.105***	9.135***	11.38***	11.38***
	(0.0960)	(0.0943)	(0.166)	(0.166)
trade route hit by SARS=1	-0.0908***	-0.0683**	-0.0389	-0.0604
	(0.0262)	(0.0288)	(0.0478)	(0.0468)
trade route hit by SARS=1 x pre SARS sourcing intensity		-0.396***		0.0727
		(0.134)		(0.378)
other routes hit by SARS=1		-0.00554		0.0540
		(0.0222)		(0.0438)
firm, destination-time, ownership, industry, origin, customs FE	Y	Y	Y	Y
R^2	0.483	0.483	0.439	0.439
No. of observations	1572882	1572882	446827	446827

Notes: This table examines the resilience of importers which has only single import/export location (roughly county level unit) in the customs data v.s. those have multiple. Column (1) and (2) include importers that only reports one single location. Column (3) and (4) include importers that reports multiple. The numbers in the parentheses are standard errors clustered at region-industry level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table A10: List of Chinese Customs Districts

Customs ID	Customs Name	Province	largest gateway	2nd largest gateway	3rd largest gateway	oversea Chinese share
100	Beijing	Beijing	Beijing			0.49%
200	Tianjin	Tianjin	Tianjin			0.01%
400	Shijiazhuang	Hebei	Tangshan	Qinhuangdao		0.00%
500	Taiyuan	Shanxi	Taiyuan			0.00%
600	Manchuri	Inner Mongolia	Manchuri			0.00%
700	Mongolia	Inner Mongolia	Baotou			0.00%
800	Shenyang	Liaoning	Shenyang			0.00%
900	Dalian	Liaoning	Dalian	Yinkou	Dandong	0.04%
1500	Changchun	Jilin	Changchun			0.00%
1900	Harbin	Heilongjiang	Harbin			0.00%
2200	Shanghai	Shanghai	Shanghai			0.18%
2300	Nanjing	Jiangsu	Suzhou	Nanjing	Lianyungang	0.04%
2900	Hangzhou	Zhejiang	Jiaxing			0.08%
3100	Ningbo	Zhejiang	Ningbo-Zhoushan			0.07%
3300	Hefei	Anhui	Wuhu			0.01%
3500	Fuzhou	Fujian	Fuzhou			0.20%
3700	Xiamen	Fujian	Xiamen	Quanzhou		1.14%
4000	Nanchang	Jiangxi	Nanchang			0.01%
4200	Qingdao	Shandong	Qingdao	Rizhao	Yantai	0.01%
4600	Zhengzhou	Henan	Zhengzhou			0.01%
4700	Wuhan	Hubei	Wuhan			0.01%
4900	Changsha	Hunan	Changsha			0.02%
5100	Guangzhou	Guangdong	Huangpu			2.75%
5200	Huangpu	Guangdong	Humen			0.01%
5300	Shenzhen	Guangdong	Shenzhen			2.07%
5700	Gongbei	Guangdong	Zhuhai			1.19%
6000	Shantou	Guangdong	Shantou			0.93%
6400	Haikou	Hainan	Haikou			0.27%
6700	Zhanjiang	Guangdong	Zhanjiang			0.01%
6800	Jiangmen	Guangdong	Jiangmen			1.02%
7200	Nanning	Guangxi	Fanchenggang			0.15%
7900	Chengdu	Sichuan	Chengdu			0.01%
8000	Chongqing	Chongqing	Chongqing			0.02%
8300	Guiyang	Guizhou	Guiyang			0.04%
8600	Kunming	Yunnan	Kunming			0.01%
8800	Lasa	Tibet	Lasa			0.00%
9000	Xi'an	Shannxi	Xi'an			0.29%
9400	Wulumuqi	Xinjiang	Wulumuqi			0.01%
9500	Lanzhou	Gansu	Lanzhou			0.00%
9600	Yinchuan	Ningxia	Yinchuan			0.01%
9700	Xining	Qinghai	Xining			0.00%

Notes: The table lists the customs district as shown in Figure A2. It also lists the gateway city (cities) for each customs district. The column on overseas Chinese is constructed from Chinese City Yearbook 1995 and defined as the number of overseas Chinese divided by local population.

Table A11: Trade Elasticity: Robustness θ

Dependent Variable: foreign sourcing relative to home sourcing $ln(\chi_{ijk}) - ln\frac{\phi^d}{\Phi}$	(1)	(2)	(3)	(4)	(5)
Weighted tariff, no log approximation	-5.180***				
	(0.748)				
lagged share weighted ln(tariff)		-5.087***			
		(0.759)			
current share weighted ln(tariff)			-5.121***		
			(0.770)		
FH firm-market import tariff				-4.989***	-5.676***
				(0.800)	(0.830)
ln maritime distance				-0.331***	
				(0.0145)	
ln geodist customs district-destination	-0.0522***	-0.0521***	-0.0522***	-0.253***	
	(0.0159)	(0.0159)	(0.0159)	(0.0561)	
ln geodist origin-customs district	-0.411***	-0.411***	-0.411***		
	(0.0153)	(0.0153)	(0.0153)		
common custom district	0.446***	0.446***	0.446***	0.411***	
	(0.0673)	(0.0673)	(0.0673)	(0.0799)	
common language: customs district-destination	0.845***	0.844***	0.844***	0.426***	
	(0.0880)	(0.0880)	(0.0881)	(0.119)	
co-Chinese	10.48***	10.48***	10.48***	8.047***	
	(2.627)	(2.627)	(2.627)	(2.741)	
Firm FE	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y
Ownership type FE	Y	Y	Y	Y	Y
Customs district FE	Y	Y	Y	Y	N
Region FE	Y	Y	Y	Y	N
Origin-Customs district-Region FE	Y	Y	Y	Y	N
R^2	N	N	N	N	Y
No. of observations	0.458	0.458	0.458	0.460	0.504
N	121742	121742	121742	114732	115964

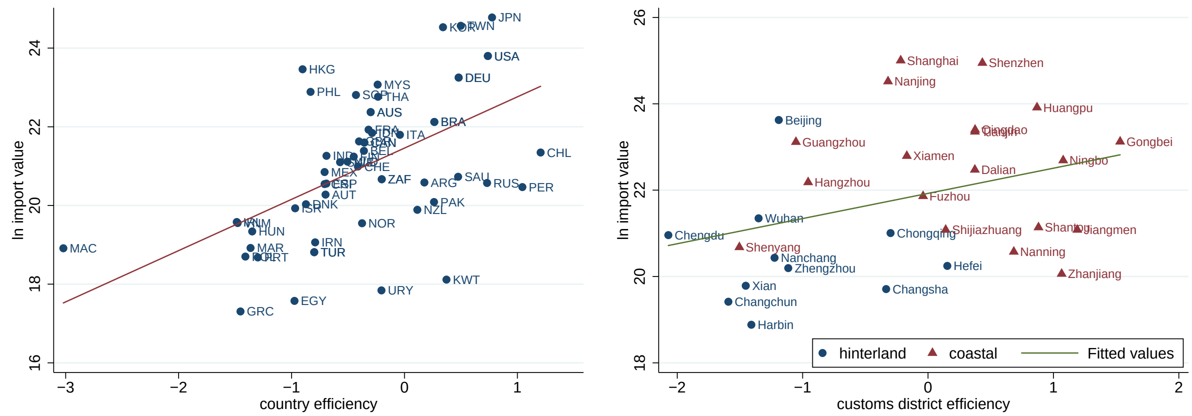
Notes: This table examines the robustness of θ with alternative measures and specifications. Column (1) uses the simple weighted average of tariff without log approximation as in Fitzgerald and Haller (2014). Column (2) uses only lagged shares in constructing the firm-market specific import tariff. Column (3) uses only current shares in constructing the tariff measure. Column (4) uses maritime distances between major ports instead of great circle to measure distances between customs districts and origins. Column (5) absorbs the gravity variables by a origin-custom district-destination fixed effect. The numbers in the parentheses are standard errors clustered at firm level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table A12: SARS cases by country

Areas	Female	Male	Total	Number of deaths	fatality ratio (%)	Date onset first probable case	Date onset last probable case
China	2674	2607	5327	349	7	16-Nov-02	03-Jun-03
China, Hong Kong	977	778	1755	299	17	15-Feb-03	31-May-03
China, Taiwan	218	128	346	37	11	25-Feb-03	15-Jun-03
Canada	151	100	251	43	17	23-Feb-03	12-Jun-03
Singapore	161	77	238	33	14	25-Feb-03	05-May-03
Vietnam	39	24	63	5	8	23-Feb-03	14-Apr-03
United States	14	15	29	0	0	24-Feb-03	13-Jul-03
Philippines	8	6	14	2	14	25-Feb-03	05-May-03
Germany	4	5	9	0	0	09-Mar-03	06-May-03
Mongolia	8	1	9	0	0	31-Mar-03	06-May-03
Thailand	5	4	9	2	22	11-Mar-03	27-May-03
France	1	6	7	1	14	21-Mar-03	03-May-03
Australia	4	2	6	0	0	26-Feb-03	01-Apr-03
Malaysia	1	4	5	2	40	14-Mar-03	22-Apr-03
Sweden	3	2	5	0	0	28-Mar-03	23-Apr-03
Italy	1	3	4	0	0	12-Mar-03	20-Apr-03
United Kingdom	2	2	4	0	0	01-Mar-03	01-Apr-03
India	0	3	3	0	0	25-Apr-03	06-May-03
Republic of Korea	0	3	3	0	0	25-Apr-03	10-May-03
Indonesia	0	2	2	0	0	06-Apr-03	17-Apr-03
China, Macao	0	1	1	0	0	05-May-03	05-May-03
Kuwait	1	0	1	0	0	09-Apr-03	09-Apr-03
New Zealand	1	0	1	0	0	20-Apr-03	20-Apr-03
Republic of Ireland	0	1	1	0	0	27-Feb-03	27-Feb-03
Romania	0	1	1	0	0	19-Mar-03	19-Mar-03
Russian Federation	0	1	1	0	0	05-May-03	05-May-03
South Africa	0	1	1	1	100	03-Apr-03	03-Apr-03
Spain	0	1	1	0	0	26-Mar-03	26-Mar-03
Switzerland	0	1	1	0	0	09-Mar-03	09-Mar-03
Total	4273	3779	8098	774	9.6		

Notes: Source: WHO.

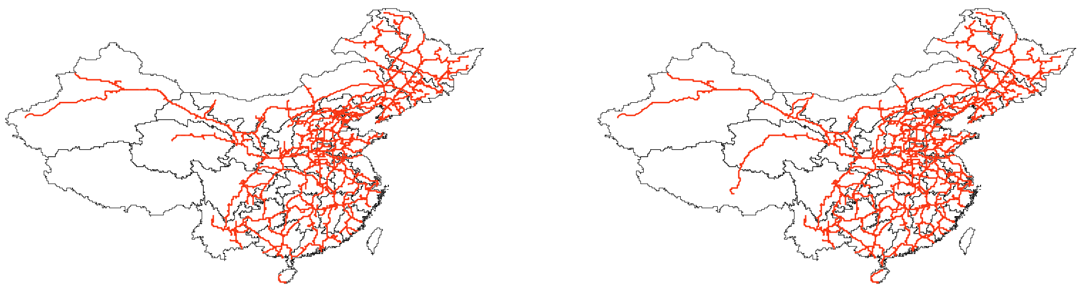
9.2 Additional Figures



(a) countries (b) customs districts

Figure A1: Country and customs district efficiencies v.s. imports

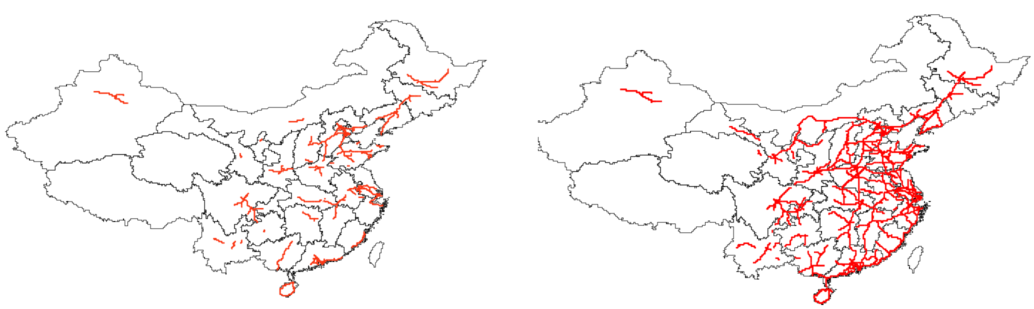
Figure A2: Map of Chinese Customs Districts



(a): year 2000 (b): year 2006

Notes: The base map is from the ACASIAN Data Center with finishing date cross-validated using news reports, government reports and other online sources.

Figure A3: Chinese Railway 2000-2006



(a): year 2000 (b): year 2006

Notes: The base map is from the ACASIAN Data Center with finishing date cross-validated using news reports, government reports and other online sources.

Figure A4: Chinese Highway 2000-2006