Production Networks and Firm-level Elasticities of Substitution*

Brian Cevallos Fujiy† Devaki Ghose‡ Gaurav Khanna§

This version: April 2023

Abstract

We provide one of the first estimates of elasticities of substitution across suppliers within the same product. We estimate these elasticities by using new real-time administrative tax data on product-level prices and quantities with firm-to-firm transactions, and leveraging the geographic and temporal variation from the Covid-19 lockdowns in India. Suppliers are highly complementary even at this granular level, with an estimated elasticity of 0.55, thus amplifying negative shocks by transmitting them through the supply chain. We quantify this transmission and show that under our estimated elasticities, the overall fall in output is substantial and widespread. In policy counterfactuals, we quantify the importance of firm connectivity separately from firm size, and of targeting aid to connected firms. Protecting more connected firms mitigates output declines non-linearly with the size of the productivity shock.

Keywords: production networks, elasticities of substitution, shock propagation, resilience
JEL Codes: D57, E32, E61, F10, L14

*We thank participants at the Danish International Economics Workshop, World Bank, CEPR's Globalisation, Strategic Autonomy, and Economic Resilience in Time of Crisis, Annual Mid-Atlantic International Trade Workshop, Midwest International Economic Development Conference, STEG's Annual Conference, IMF's Jacques Polak Annual Research Conference, Bank of Italy's Global Value Chains, George Washington University, and Lund University for their comments and suggestions. We thank David Baqaee, Ana Cecilia Fieler, Banu Demir, Pablo Garriga, Andrei Levchenko, Sebnem Kalemli-Ozcan, Scott Orr, Ludovic Panon, Carolina Sanchez, Ritam Chaurey, and Sebastian Sotelo for insightful discussions. We thank Ahmed Magdy and Eli Mogel for their excellent research assistance and Sushil Khanna for guidance on the data and details on the analysis. The views expressed in this paper do not represent the views of the World Bank or its partner organizations and solely represent the authors' personal views.

†University of Michigan, bcfujiy@umich.edu.
‡The World Bank, dghose@worldbank.org.
§University of California, San Diego, gakhanna@ucsd.edu.

1 Introduction

The ability of firms to substitute inputs across suppliers is critical for the resilience of supply chains and the transmission of supply shocks. If it is difficult for firms to substitute across suppliers after an adverse supply shock, the shock could amplify by transmitting further downstream and upstream through the supply chain. The importance of this mechanism was reflected during the Covid-19 pandemic, where supply chain disruptions drove dramatic reductions in GDP worldwide. For instance, India reported a -7.3% growth rate for the 2020/21 financial year, one of the most significant contractions worldwide and the largest decline in GDP since India's independence.[1]

In this paper, we quantify the importance of firm-level elasticities of substitution across suppliers (firms) of the same product category to explain large fluctuations in GDP. We provide new estimation strategies and estimates for these elasticities by leveraging regional variation in supply-side shocks induced by the Indian government's massive lockdown policy. We show that this elasticity is key to partly explaining the dramatic decline in incomes during the Covid-19 pandemic. Using new big data computational techniques, we quantify this decline in GDP by directly leveraging information on the economy-wide firm-to-firm network, and highlight how protecting more connected firms mitigates output declines.

We pose two main research questions. First, are suppliers of intermediate inputs within a product category complements or substitutes? The answer to this question determines how shocks propagate throughout supply chains. We expect shocks to propagate less across firm networks if input suppliers are substitutable. However, if input suppliers are complements, the effects of adverse shocks can easily propagate through buyer-supplier networks. Second, we ask, how this newly estimated elasticity affects firm-level sales, and ultimately GDP, by propagating and amplifying negative shocks through firm-level input-output linkages, and how the effect of the shock depends on the connectivity of affected firms.

Two unique features of our setting allow us to answer these questions credibly. First, India had a distinct mosaic of lockdown policies, whereby the roughly 600 districts were classified into three different zones with varying degrees of restrictions. This allows us to isolate variation in the ability to trade and transport goods over this period. Second, we obtain new granular and high-frequency administrative data on the universe of firm-to-firm transactions for a state in India, with unique information on unit values and HS-product classifications. These data, while not used before, allow us to estimate new elasticities at the firm (rather than industry) level and across

[1] https://www.economicsobservatory.com/how-has-Covid-19-affected-indias-economy. More broadly, during the 2020/21 financial year, GDP fell by -3.3% in emerging market economies, and by -2.2% in developing countries.

different suppliers of a product (rather than across products).

We begin by presenting reduced-form evidence that leverages the nationwide, sudden, and unprecedented lockdown imposed by the Indian government in March 2020. Importantly, these lockdowns induced geographic variation: districts were categorized into *Green* (mild lockdown), *Orange* (medium lockdown), and *Red* (severe lockdown). Since the lockdowns were sudden and unexpected, they were likely implemented independent of economic fundamentals and induced strong variation in transactions between firms across India.[2] We show, conditional on high-dimensional fixed effects, that these adverse supply shocks led to a sharp increase in unit values (prices), and dramatic fall in transactions of intermediate inputs, if either buyers or sellers were located in high lockdown zones.

Next, we modify a standard multi-sector firm-level model of input-output linkages by augmenting the production function with substitution across suppliers within the same product category. We derive analytical expressions that relate the relative values of quantities purchased of the same product from different suppliers to the equilibrium relative prices. A combination of the lockdown-induced price variation, and this analytical framework allows us to estimate how substitutable the different suppliers are, within each product category.

Yet, Covid-19 was not just a supply shock. Baqaee and Farhi (2020) argue that the pandemic outbreak was a combination of exogenous shocks to the quantities of factors supplied, the productivity of producers, and the composition of final demand by consumers across industries. To estimate the elasticity of substitution across suppliers of inputs, we leverage variation in input prices driven by the sudden restrictions in economic activity in lockdown districts where these suppliers were located. In addition, we leverage variation in trade costs arising from transportation restrictions in districts via which the goods need to pass from the seller to the buyer. While our instruments help derive the necessary variation, to further isolate supply shocks from other shocks, we control for an entire array of high-dimensional fixed effects, such as product-by-month fixed effects to account for product-level shocks, and buyer-by-month fixed effects to account for demand-side shocks. Given the richness of our product data, we can also include buyer-by-product and seller-by-product fixed effects. We further control for other factors, such as firms' exposure to foreign shocks transmitted through trade (Hummels et al., 2014), and the caseload and severity of Covid-19 cases.

We find that suppliers within the same HS-4 product category are highly complementary. Our estimated elasticity of substitution across suppliers of the same product is 0.55. In various specification tests employing different combinations of fixed effects and different sources of vari-

[2] https://www.bbc.com/news/world-asia-india-56561095, https://thewire.in/government/india-Covid-19-lockdown-failure

ation, we find that the estimated elasticities lie within a range of 0.50 to 0.66. Our new elasticities show that inputs across firms are highly complementary even within the same HS-4 product category. The elasticities are similar at the HS-6 product level and even smaller at the HS-8 product level.[3] As a result, second-order effects amplify adverse firm-level shocks by propagating through the network, and contributing to GDP fluctuations. Additionally, we also estimate a more aggregate elasticity of substitution across different industries of 0.69, which implies complementarity across industries as in Atalay (2017) and Boehm et al. (2019).

We provide two additional empirical insights. First, the elasticity of substitution varies across industries and locations, and exhibits more complementarity in districts with better contract enforcement and institutional quality. This reflects the possibility that products or industries that are relationship-specific exhibit more complementarity, and benefit more from better contract enforcement. Second, given the expectations that the shock may be short-lived, our estimates are relevant for short-run shocks rather than longer-term structural changes. We show that as we increase the time-frequency of adjustments, the elasticities increase but still display complementarity.

The literature so far provides little guidance about estimates of the firm-level elasticity of substitution between suppliers within product categories (Taschereau-Dumouchel, 2020; Baqaee and Farhi, 2019; Jones, 2011). While other work estimates elasticities of substitution across industries (Atalay, 2017), across products from different countries (Boehm et al., 2019), or across intermediate goods (Carvalho et al., 2021; Peter and Ruane, 2022), such estimates do not yet exist for substitution elasticities across suppliers within the same product category. Estimating elasticities of substitution across different suppliers has been challenging for two reasons. First, it is difficult to find detailed information on product-specific unit values reported by each firm in firm-to-firm transactions. Second, it is challenging to find exogenous sources of variation in firm-level prices (rather than product-level prices) that allow one to estimate these elasticities credibly.

In the final part of our paper, we embed the estimated elasticities in our model and analyze how input complementarities at the firm level affect aggregate economic outcomes. Considering a firm's full network is a highly computationally demanding task: In this case, it involves inverting a 94,555 by 94,555 input-output matrix.[4] We apply state-of-the-art computational techniques from computer science to our input-output matrix to derive the full connectivity of each firm.

Compared to our baseline case ($\epsilon = 0.55$), we find that the quarterly fall in GDP induced by a negative 25% shock to firms in *Red* zones would be $2.68pp$ less in a model where firms in the same HS-4 product category are substitutes ($\epsilon = 2$) and $0.99pp$ more when they are almost Leontief

[3] We use the term "product" and "product category" interchangeably and define whether we refer to HS-4, HS-6, or HS-8 codes when relevant. We use the term "industry" to refer to the broad HS Sections.

[4] As a reference, the typical sector-level input-output matrix from the BEA 2012 is 405 by 405.

($\epsilon = 0.001$).[5] The additional losses due to firm-level complementarities translate into 870 million USD, which is about 25 USD per capita per quarter. Additionally, under our estimated elasticity, the shock to *Red* zone firms spreads widely through the network to other parts of the economy.

In our policy counterfactuals, we quantify the importance of firm connectivity separately from firm size and the effects of bailing out the most connected firms (measured by all direct and indirect connections) instead of bailing out only the directly connected firms. In policy and academic circles, much importance has been paid to large firms, as Hulten (1978) emphasized the importance of firm sizes in the propagation of shocks through production networks. We show that controlling for size, the fall in GDP is much larger if the most connected firms are affected compared to the least connected or a random set of firms. The importance of the most connected firms increases non-linearly with the size of the negative productivity shock and decreases as firms become more substitutable. Our experiment suggests that when comparing firms of the same size, governments should target bailouts to better-connected firms.

Finally, we quantify how important it is to also consider a firm's indirect connectivity, and find that under our estimated elasticity and a negative productivity shock of 25%, the fall in GDP would be $2.56pp$ less if the government were to bail out firms on the basis of total connectivity as opposed to direct connectivity (counting only the number of direct buyers of a supplier).[6] We see that as the shock gets larger, the difference in aggregate GDP between these two experiments rises, emphasizing the importance of measuring a firm's indirect connections as well.

Related Work. Our paper contributes to two strands of the literature. First, we contribute to the literature on shock propagation and amplification through supply chains and production networks (Barrot and Sauvagnat, 2018; Carvalho et al., 2021; Peter and Ruane, 2022; Boehm et al., 2019; Korovkin and Makarin, 2020; Ferrari, 2022; Dew-Becker, 2022; Huneeus, 2018; Arkolakis et al., 2023). A crucial parameter determining the degree of shock propagation through supply chains is the firm-level elasticity of substitution across suppliers within the same industry. There are at least two estimation challenges highlighted by this literature. First, most firm-to-firm datasets do not contain product-level (unit) prices from each supplying firm to a buying firm or lack the required variation in such prices to estimate firm-level elasticities of substitution across suppliers.[7] Second,

[5] We find that a 25% productivity shock to firms in *Red* zones reduces GDP by 10.95%. As an empirical benchmark, the state's annual GDP fell by 11.3% in 2020/21.

[6] A firm's indirect connections measure not only the number of direct buyers of a supplier but also the buyers' buyers and their buyers, and so on.

[7] Carvalho et al. (2021) observe a binary measure of whether firms were connected via buyer-supplier relationships, rather than quantities and unit values associated with such transactions. They use a proportionality assumption which precludes estimating the elasticity of substitution across suppliers within a product category, as a buyer sourcing from two suppliers in the same industry will source the same amount given the assumption. Although lacking firm-to-firm price data, Dhyne et al. (2022) structurally estimate a similar elasticity in the context of imperfect competition models

existing work has estimated firm-level elasticities of substitution between product categories or between domestic and foreign industries (Peter and Ruane, 2022; Carvalho et al., 2021; Boehm et al., 2019; Atalay, 2017), but limited identifying variation in prices at the buyer-supplier level has not allowed one to estimate this elasticity across suppliers within a product category. We provide one of the first estimates of the elasticity of substitution across suppliers within a product category. The lack of firm-level elasticities across suppliers has so far constrained our assessment of the importance of nodal firms, such as the most connected firms, in the propagation of shocks through production networks.

We contribute to the literature in each of these dimensions. First, we measure unit values (prices) and quantities at the seller-buyer-product-transaction level. We derive price changes from supply and transportation disruptions in lockdown-affected districts and estimate the firm-level elasticity of substitution between suppliers within a product category. We then quantify how this elasticity amplifies firm-specific supply shocks through a roundabout production network (Baqaee and Farhi, 2019). Finally, given the presence of complementarity across suppliers of the same product, we address previously unanswered questions on the importance of firms' overall connectivity within a production network for the amplification of shocks. We leverage computational innovations in big data to compute the second-order effects of productivity shocks using the entire matrix of production linkages. This innovation helps quantify the non-linear effects of productivity shocks *directly* using the network, without relying on approximations using final sales.[8]

Our paper is also related to research on trade collapses during adverse shocks (Behrens et al., 2013; Giovanni and Levchenko, 2009; Bricongne et al., 2012), and shock transmission through GVCs during the Covid-19 pandemic via disruptions to imports, exports, or aggregate production (Bonadio et al., 2021; Baqaee and Farhi, 2020; Cakmakli et al., 2021; Demir and Javorcik, 2020; Gerschel et al., 2020; Heise, 2020; Lafrogne-Joussier et al., 2022; Bas et al., 2023; Chakrabati et al., 2021; Khanna et al., 2022; Miranda-Pinto et al., 2022). In contrast, we analyze how domestic transactions were affected during Covid lockdowns in a large developing country. Our policy motivation stems from the observation that policymakers worldwide are interested in quantifying the trade-off between strict lockdowns that affect GDP through complex buyer-seller networks, and more lenient measures that increase production and trade. More importantly, even beyond the immediate crises, our estimates of how substitutable suppliers are within a product category will help policymakers quantify the economy-wide effects of any disruptive events (e.g., natural disasters or sanctions) on trade and production that are expected to be reasonably short lived.[9]

where they restrict the elasticity to be larger than 1 for mark-ups to be relevant.

[8]As firm-to-firm data become common (Panigrahi, 2021; Demir et al., 2021; Dhyne et al., 2021; Alfaro-Urena et al., 2020; Chacha et al., 2022), our methods can be used to quantify shock propagation through complex networks.

[9]We may hesitate to use these elasticities for exercises on long-term structural transformations.

2 Data and Context

Firm-to-firm trade. Our primary data source is daily establishment-level transactions with distinct information on establishment locations.[10] This data is provided by the tax authority of a large Indian state with a diversified production structure, around 50% urbanization rates, and high levels of population density. To benchmark the size of this Indian state to other firm-to-firm trade datasets, the population of this state is roughly three times the population of Belgium, seven times the population of Costa Rica, and double the population of Chile.

The data contains daily transactions between all registered establishments in this state and all registered establishments in India and abroad, from April 2018 to October 2020. This data is collected by the tax authority's *E-way Bill* system to increase compliance for tax purposes. This is an advantage over standard VAT firm-to-firm trade datasets in developing countries, which suffer from severe under-reporting. By law, anyone dealing with the supply of goods and services whose transaction value exceeds Rs 50,000 (700 USD) must generate E-way bills. Transactions with values lower than 700 USD can also be registered, but it is not mandatory. The E-way bill is generated before transport (usually via truck, rail, air, or ship), and the vehicle driver must carry the bill with them, or the entire extent of goods can be confiscated. Our data is generated from these bills. This implies that our network is representative of relatively larger firms, but the threshold is sufficiently low that we are likely capturing small firms as well.

Each transaction reports a unique tax code identifier for both the selling and buying establishments, all the items contained within the transaction, the value of the whole transaction, the value of the items being traded up to 8-digit HS codes,[11] quantity of each item, units, and mode of transportation. Each transaction also reports the ZIP code of both the selling and buying firms, which we use to merge with other district-level datasets.

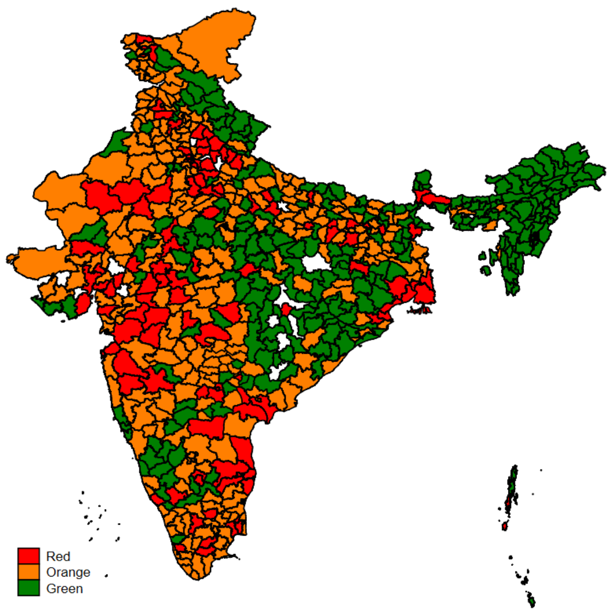
Since the data report both value and quantity of traded items, we construct unit values for each transaction. We also calculate average unit values at the 4-digit product-by-month-by-seller-by-buyer level, the number of transactions, and the total value of the goods transacted. This is the foundation of our firm-to-firm dataset used in the analysis.

[10] While we use the term "firm" throughout the paper, our data is actually at the more granular establishment level, and we can identify the parent firms for each establishment as well.

[11] The data partially reports items up to 8-digit HS codes. Until April 2021, in India, it was only mandatory to report 4-digit HS codes of goods traded. See https://economictimes.indiatimes.com/small-biz/gst/six-digit-hsn-code-in-gst-made-mandatory-from-april-1/articleshow/81780235.cms?from=mdr. 97% of transactions report 4-digit HS codes, and 40% report 8-digit HS codes. Given this, our main specifications are based on 4-digit HS codes, and we show robustness to using HS-8 codes.

Lockdowns. On March 25th, 2020, India unexpectedly imposed strict lockdown policies nationwide. Districts were classified into *Red*, *Orange*, and *Green* zones according to each district's severity of Covid cases. Each color corresponds to different lockdown degrees, where *Red* was severe lockdown, *Orange* was medium lockdown, and *Green* was mild. Yet, at that time, there were barely any Covid cases in India, as the entire country averaged about 50 cases a day (as opposed to about 400,000 cases a day the following year).

FIGURE 1: India's Covid-19 lockdown zones



Notes: The map shows the lockdown zones across Indian districts announced on March 25, 2020.

In Figure 1, we map the distribution of lockdowns across India. Districts in the *Red* zone experienced the strictest lockdowns, where rickshaws, taxis and cabs, public transport, barbers, spas, and salons remained shut. E-commerce was allowed for essential services. *Orange* and *Green* zone districts experienced fewer restrictions. *Orange* zones allowed the operation of taxis and cab aggregators, as well as the inter-district movement of individuals and vehicles for permitted activities. In addition to the activities allowed in *Orange* zones, buses were allowed to operate with up to 50% seating capacity and bus depots with 50% capacity in *Green* zones.[12]

Throughout the paper, we use this color scheme as the treatment across Indian districts. In particular, each firm is located within a district, so treated firms are located within a *Red*, *Orange*,

[12] https://economictimes.indiatimes.com/news/politics-and-nation/lockdown-3-0-guidelines-for-red-zone/activities-prohibited/slideshow/75503925.cms. On April 30, one *Red* zone district was reclassified to the *Green* zone, but we maintain the initial classification as it is likely to be more exogenous.

or *Green* district between March and May 2020.

Physical and cultural distance. We use different measures of *distance*, which we include as controls in our empirical results. The measures of geographic distance between districts calculate the length of the shortest distance between district centroids. The measure of linguistic distance between Indian districts is from Kone et al. (2018) using the commonly used ethnolinguistic fractionalization (EFL) index (Mira, 1964). This index measures the probability of two randomly chosen individuals from different districts speaking the same language.

Other controls. We control for different firm and district-level time-varying variables such as data on the monthly number of cases, deaths, and recoveries from Covid-19 for all of India at the district level from `www.Covidindia.org`. For each firm, we construct two variables that measure the firm's exposure to global demand and supply shocks that vary at the product and country level, following Hummels et al. (2014). The construction of these exposure variables is described in detail in online data Appendix C.

Summary statistics. We present some key summary statistics from our firm-to-firm trade dataset in Table A1. Panels A and B report the unique numbers of sellers and buyers, total sales (in million rupees), and total number of transactions during the months of January-March, April-June, and July-September, for both 2019 and 2020. The most noticeable pattern is the large drop in all variables in 2020 compared to 2019, particularly during the April-June period, which coincided with the lockdown policies. The total value of sales and the number of transactions fell by almost 60% during April-June of 2020 compared to 2019. For reference, the fall in the value of sales was only 25% after the strict centralized lockdown was over (July-September) and only 15.6% before the lockdown (January-March) compared to the corresponding months in 2019.

To further understand the composition of economic activity of the Indian state of our analysis, in Table A2, we show the types of goods firms sell and buy, and what fraction crosses state and country borders. In our state, firms are mostly in the business of selling vegetables, plastics, and minerals; and of buying machinery, metals, and vegetables. In terms of the type of trade, firms in our state are more likely to sell to others firms in the state. This contrasts with how firms in our state buy goods, where the share of purchases that come from within the state is almost the same as from other Indian states. Finally, international exports and imports represent a non-negligible but small share of sales and purchases.

Before using the lockdown variation to understand how firm-to-firm transactions are affected, we verify the stringency of these lockdowns in Figure A3 using Google mobility data. The data shows how the number of visitors to (or the time spent in) categorized places changes in comparison to baseline days. The baseline day is the median value from the 5-week period Jan 3 – Feb 6, 2020.[13] Until the beginning of March 2020, there were essentially no differences in mobility trends across *Red*, *Orange*, or *Green* zones. But starting at the end of March 2020, we see that there is a substantial reduction in different types of activities (time spent in retail and recreation, grocery and pharmacy, parks, commuting, and workplaces) in *Red* zones compared to *Green* zones; with *Orange* zones in between. People in *Red* zones also spent more time at home than people in either *Orange* or *Green* zones. We notice that starting in August 2020, a few months after the centralized lockdown was over, these differences reduced, and by December 2020, these differences, especially in workplace mobility, become negligible.

3 Reduced-form evidence

In this section, we outline a simple empirical specification to provide evidence showing the role of lockdown policies on key outcome variables for firm-to-firm trade. We show that the sudden Covid-19 lockdown policies between March and May 2020 led to a rise in unit values, and a fall in the monthly number of transactions between firms.[14] In the next section, we exploit this variation to estimate firm-level elasticities of substitution across intermediate suppliers of the same product.

3.1 Empirical specifications

Our reduced-form specifications implement a difference-in-differences approach where we compare the unit values and the number of transactions both at the seller and seller-buyer level across *Red*, *Orange*, and *Green* districts, before and after the lockdown. In our analysis at the seller level, the omitted (control) group are sellers located in *Green* districts, and the base month is February 2020, the month before the lockdown enforcement. At the seller-buyer level, the omitted groups are sellers and buyers located in *Green* zones, and the base month is February 2020.

Seller-level regressions. We estimate the following specification:

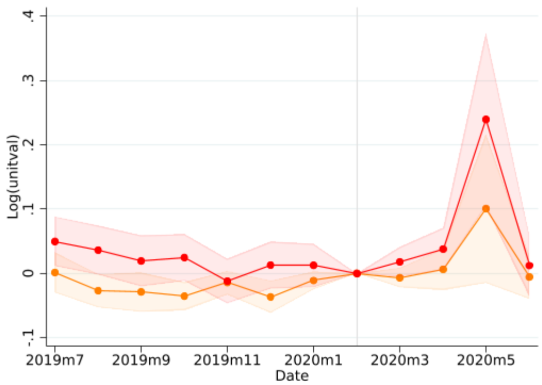
$$Y_{si,t} = \iota_{i,o_{(s)}} + \iota_{i,t} + \sum_{t \neq -1} \beta_t Red_{o_{(s)}} + \sum_{t \neq -1} \gamma_t Orange_{o_{(s)}} + \mathbf{X}\boldsymbol{\delta} + \epsilon_{si,t}, \tag{1}$$

[13] Source: https://support.google.com/covid19-mobility/answer/9824897?hl=en&ref_topic=9822927

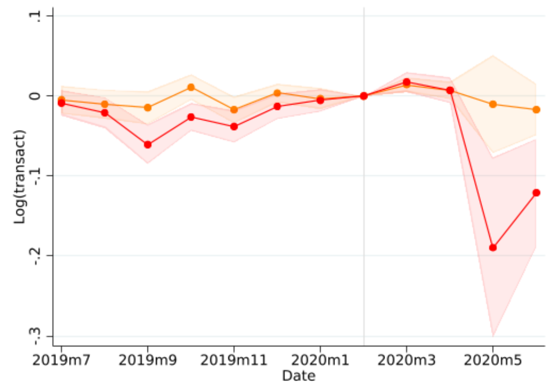
[14] To see a similar application of this empirical strategy for domestic violence and economic activity in India, see Ravindran and Shah (2020) and Beyer et al. (2021).

FIGURE 2: Seller-level reduced-form event studies

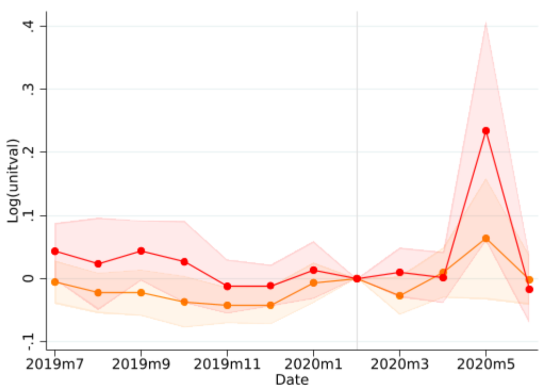
(a) Unit value, 4-digit HS



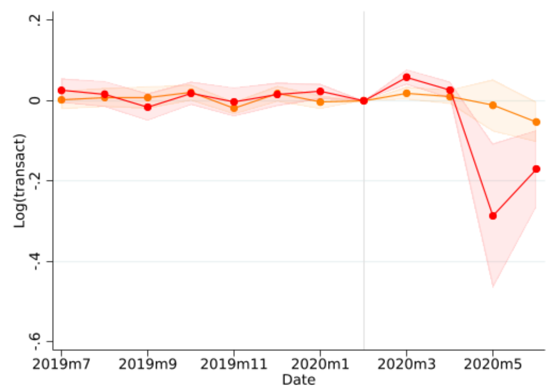
(b) # Transactions, 4-digit HS



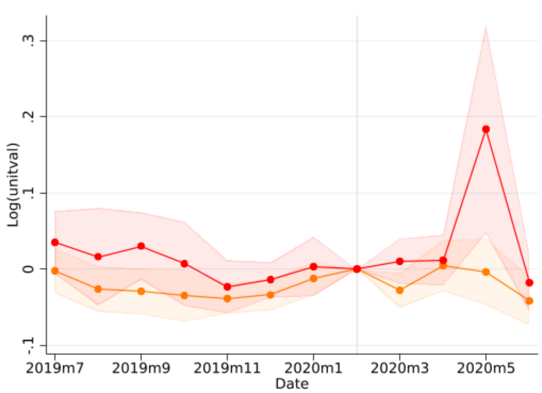
(c) Unit value, 8-digit HS



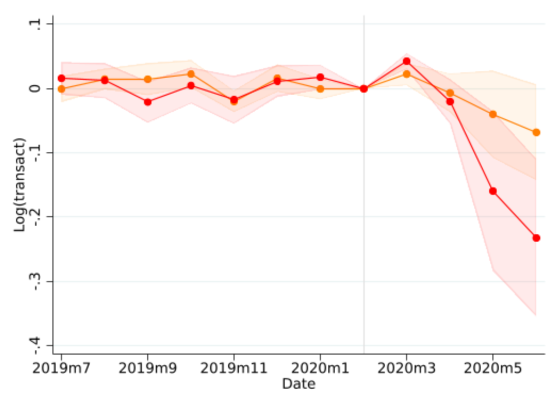
(d) # Transactions, 8-digit HS



(e) Unit value, 8-digit HS, strong FEs



(f) # Transactions, 8-digit HS, strong FEs



Notes: This figure consists of 6 plots. Each plot shows estimates for β_t and γ_t from Equation (1). The estimated values are all in comparison to sellers in *Green* districts in February 2020. The dependent variable for the plots on the left-hand side is log unit values. The dependent variable for the plots on the right-hand side is the log number of transactions. Each row varies by the definition of a product-group, and the fixed effects included in the regression. In the first row, a product is a 4-digit HS code, and the fixed effects are product-by-month and district. In the second row, a product is an 8-digit HS code, and the fixed effects are product-by-month and district. In the third row, a product is an 8-digit HS code, and the fixed effects are product-by-month and district-by-product. Standard errors are clustered at the district level. All controls mentioned in the paper are included. The shaded areas are 95% confidence intervals.

where $Y_{si,t}$ are either unit values or the log number of transactions for seller s of HS-4 product i in month t, $\iota_{i,t}$ are product-by-month fixed effects, $\iota_{i,o_{(s)}}$ are product-by-district fixed effects (i.e. fixed effects based on the district o where seller s resides). $\mathbf{X}$ are controls that include the number of Covid cases, deaths, recoveries, and exposure to international demand and supply shocks as discussed in Appendix C. We control for the Covid cases, and deaths since these are the variables on which the government based its lockdown decisions (Ravindran and Shah, 2020). The covariates of interest are $Red_{o_{(s)}}$ and $Orange_{o_{(s)}}$. The first is an indicator variable that equals 1 if seller s located in district $o_{(s)}$ experienced a severe lockdown, 0 otherwise. The second equals 1 if seller s located in district $o_{(s)}$ experienced a mid-level lockdown, 0 otherwise. The excluded category is $Green_o$ districts, where mild lockdown was imposed. The estimates of interest are β_t and γ_t. Our base time category is February 2020, which is just before lockdowns began. Standard errors are clustered at the seller's origin district level.

Seller-buyer level regressions. At the seller-buyer level, we estimate the specification:

$$Y_{si,b,t} = \sum_{(v,z)\in\Omega} \sum_{t\neq -1} \beta_t^{vz} \left(\gamma_{o_{(s)}}^{v} \times \gamma_{d_{(b)}}^{z}\right) + \delta_{o_{(s)}} + \delta_{d_{(b)}} + \delta_{i,t} + \beta_1 \log\ dist_{od} + \mathbf{X}\boldsymbol{\delta} + \epsilon_{si,b,t}\ , \tag{2}$$

where $Y_{si,b,t}$ are unit values or the number of transactions in logs between seller s of HS-4 product i and a buyer b in month t. $\delta_{o_{(s)}}$, $\delta_{d_{(b)}}$, and $\delta_{i,t}$ are origin, destination, product-by-month fixed effects. $dist_{od}$ is a vector of cultural and geographic distance variables, and $\mathbf{X}$ are controls that include the number of Covid-19 cases, deaths, recoveries, and exposures to international demand and supply shocks. The first term of the right-hand side contains our estimates of interest. $(v,z)\in\Omega$ is a duple that contains the color y of seller's district, and the color z of buyer's district. Ω is the set that includes all pairs except $(Green, Green)$, such that this is the excluded category when estimating Equation (2). $\gamma_{o_{(s)}}^{v}$ and $\gamma_{d_{(b)}}^{z}$ are thus indicator variables that equal 1 when seller s is located in district o located in lockdown zone v, and when buyer b is located in district d located in lockdown zone z, respectively. The estimates of interest are β_t^{vz}. Our base time category is February 2020, which is just before lockdowns began. Standard errors are two-way clustered at the origin and destination district levels.

3.2 Reduced-Form Facts

Fact 1: Sellers' unit values disproportionately rose, and trade fell in more severe lockdown zones. The first two panels of Figure 2 plot the coefficients β_t and γ_t from Equation (1), representing changes in log unit values and log number of transactions with respect to *Green* districts in

February 2020 (the base category). In May 2020, sellers' unit values in *Red* districts rose by $25pp$, and in *Orange* districts rose by around $10pp$ with respect to the base category.

At the same time, sellers' number of transactions in *Red* districts declined by around $20pp$, and in *Orange* districts declined by around $3pp$ with respect to the base category. Additionally, as expected by the severity of the lockdown policies by color, the rise in unit values and fall in the number of transactions was larger for sellers in *Red* districts than for *Orange* ones. In both figures, we find no evidence of pre-trends, implying that there were likely no differences in the trends of unit values or number of transactions between *Red*, *Orange*, and *Green* districts before the lockdown.

The middle two panels of Figure 2 repeat the same exercise with a finer product definition, using 8-digit HS codes. Results remain virtually the same. In the last row of Figure 2, we include a stronger set of fixed effects (e.g., district-by-product), and the results remain the same.

Fact 2: Equilibrium unit values rose, and the number of transactions fell in more severe lockdown zones. We now report the results from our seller-by-buyer-level specification. In Figures 3 and 4, we report the estimates for β_t^{yz} from Equation (2), where the estimates are in comparison to cases when both sellers and buyers were located in *Green* districts in February 2020.

In the first row of Figure 3, we plot the coefficients from regression (2) where the seller is in the *Red* zone, and the buyer is in *Red*, *Orange*, and *Green* zones respectively. Similarly, in the second row of Figure 3, we plot the coefficients from regression (2) where the seller is in the *Orange* zone, and in the third row, we plot the coefficients from regression (2) where the seller is in the *Green* zone (and the buyer is in *Red* and *Orange* zones respectively).

There are two main takeaways from these figures. First, even after controlling for bilateral resistance terms, trade costs, and additional covariates, unit values rose, and the number of transactions fell with respect to the base category (both buyer and seller in *Green* zones). The rise in unit values was as much as $45pp$, and the fall in transactions was as high as $12pp$. Second, these changes seem proportional to the severity of the lockdowns for both sellers and buyers. Once again, there is no evidence of differential pre-trends across zones leading up to the shock.

Our two facts jointly imply that prices, where either sellers or buyers were located in *Red* districts, were higher during the lockdown in comparison to districts where the lockdowns were mild (*Green* zones). As such, the lockdown induced variation in prices that we will later leverage to estimate elasticities of substitution across intermediates.

FIGURE 3: Unit Value, Seller-Buyer-Level Regressions

Notes: This figure consists of 8 plots. Each plot shows estimates for β_t^{yz} from Equation (2). The values of the estimates are all in comparison to sellers and buyers both in *Green* districts in February 2020. The vertical line in January 2020 splits the period into pre and post-lockdown periods. The dependent variable for the plots is the log unit value. A product is a 4-digit HS code. Regressions include product-by-month, origin, and destination district fixed effects. Standard errors are two-way clustered at the origin and destination district levels. All controls mentioned in the paper are included. The color of the line denotes the color of the seller's district, while the color of the shaded 95% confidence interval denotes the color of the buyer's district.

4 MODEL

We build a quantitative general equilibrium model of firm-to-firm trade (Baqaee and Farhi, 2019). The production sector is perfectly competitive.[15] We adapt the general nested CES structure to reflect the possibility that suppliers within the same product category could be substitutes or complements. Firms combine inputs in a CES fashion in each of its three tiers. In the first tier, firms

[15]We abstract from market power (Edmond et al., 2018; Alviarez et al., 2021) since the evidence from the data suggests that the market structure in this Indian state is highly competitive. The median HHI across 4-digit HS product categories is 0.1041, which implies a low level of market concentration within a product category.

FIGURE 4: Number of Transactions, Seller-Buyer Level Regressions

Notes: This figure consists of 8 plots. Each plot shows estimates for β_t^{yz} from Equation (2). The values of the estimates are all in comparison to sellers and buyers both in *Green* districts in February 2020. The vertical line in January 2020 splits the period into pre and post-lockdown periods. The dependent variable for the plots is the log number of transactions. A product is a 4-digit HS code. Regressions include product-by-month, origin, and destination district fixed effects. Standard errors are two-way clustered at the origin and destination district levels. All controls mentioned in the paper are included. The color of the line denotes the color of the seller's district, while the color of the shaded 95% confidence interval denotes the color of the buyer's district.

combine labor and aggregated intermediate inputs to produce output. In the second tier, firms combine aggregated intermediate inputs of a product category. In the third tier, firms combine suppliers within a product category. The model yields estimating equations we then use to estimate the firm-level elasticity of substitution between suppliers of the same product category.

We consider a fixed set of firms $\mathcal{F}$ and of product categories $\mathcal{I}$, where $N = |\mathcal{F}|$ is the total number of firms in the economy, N_i is the number of firms producing a good of product category i, and $I = |\mathcal{I}|$ is the number of product categories. Each firm produces according to its technology

$$y_{nj} = A_n \left(w_{nl} \left(l_n\right)^{\frac{\alpha-1}{\alpha}} + (1 - w_{nl}) \left(x_{nj}\right)^{\frac{\alpha-1}{\alpha}} \right)^{\frac{\alpha}{\alpha-1}}, \tag{3}$$

where y_{nj} is the output produced by firm n in product j, A_n is the productivity of firm n, l_n is the labor used by firm n, x_{nj} is the composite intermediate input used by firm n in product category j, α is the elasticity of substitution between labor and the composite material input, and w_{nl} is the intensity of labor in production. The composite material input, in turn, consists of inputs from the I different product categories in the economy, and is:

$$x_{nj} = \left(\sum_{i=1}^{I} w_{i,nj}^{\frac{1}{\zeta}} \left(x_{i,nj}\right)^{\frac{\zeta-1}{\zeta}} \right)^{\frac{\zeta}{\zeta-1}}, \tag{4}$$

where ζ is the firm-level elasticity of substitution across products i,[16] and $w_{i,nj}$ is the importance of inputs of product category i for firm n of product j. $x_{i,nj}$ are intermediate inputs from product i going to firm n producing product j, which are constructed as:[17]

$$x_{i,nj} = \left(\sum_{m=1}^{N_i} \mu_{mi,nj}^{\frac{1}{\epsilon}} x_{mi,nj}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}, \tag{5}$$

where $x_{mi,nj}$ are intermediate inputs from supplier m of product i sold to firm n producing product j. $\mu_{mi,nj}$ is the importance of input from supplier m of product i in the production of firm n of product j, and ϵ is the firm-level elasticity of substitution across different suppliers within the same product category. This is the key elasticity we want to estimate. The above production functions work for reproducible factors. For non-reproducible factors (in our case, labor), the production function is an endowment: $Y_f = 1$.

Product 0 represents the final consumption of the household and is given by

$$C = \left(\sum_{i}^{N} w_{0i} \left(c_i\right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{6}$$

where $\sum_i w_{oi} = 1$ and σ is the elasticity of substitution in consumption.

[16] Previous work has estimated different related versions of this elasticity, for e.g., elasticities of substitution across industries (Atalay, 2017), across goods from different countries (Boehm et al., 2019), or across product categories (Carvalho et al., 2021; Peter and Ruane, 2022).

[17] We exclude foreign intermediate goods since they are not exposed to Indian Covid-19 lockdown shocks.

Model in standard form. To write the economy in standard form as in Baqaee and Farhi (2020), we construct an input-output matrix $\widehat{\Omega}$ with dimension $2+N+I$, where 2 dimensions come from the household's consumption aggregator, and a factor of production (labor), N dimensions come from the N firms in the economy, and I dimensions come from the I product categories in the economy. We explicitly distinguish between labor and intermediate inputs since labor is non-reproducible.

Consider the vector of elasticities by $\widehat{\theta}$, where $\widehat{\theta} = (\sigma, \alpha, \zeta, \epsilon)$. Formally, a nested-CES economy in standard form is defined by the tuple $(\widehat{\Omega}, \widehat{\theta})$. The input-output matrix $\widehat{\Omega}$ of size $(2+N+I) \times (2+N+I)$ is a matrix where element (i, j) equals the value of $\Omega_{ij} = \frac{p_j x_{ij}}{p_i y_i}$, which is the expenditure share of the ith firm on inputs from the jth supplier as a share of the total revenue of firm i. Note that every supplier is a CES aggregate. The Leontief inverse is $\psi = (1-\Omega)^{-1}$. Intuitively, the $(i, j)th$ element of ψ (the Leontief inverse) measures i's total reliance on j as a supplier. That is, it captures both the direct and indirect ways through which i uses j in its production. Let us also denote the sales of producer i as a fraction of GDP by λ_i, where $\lambda_i = \frac{p_i y_i}{\sum_j^N p_j c_j}$.

The input-output covariance operator is

$$Cov_{\Omega_k}(\psi_{(i)}, \psi_{(j)}) = \sum_{l=1}^{2+N+I} \Omega_{kl}\psi_{li}\psi_{lj} - \left(\sum_{l=1}^{2+N+I} \Omega_{kl}\psi_{li}\right)\left(\sum_{l=1}^{2+N+I} \Omega_{kl}\psi_{lj}\right). \tag{7}$$

This operator measures the covariance between the ith and the jth columns of the Leontief inverse using the kth row of the input-output matrix as distribution. The second-order macroeconomic impact of microeconomic shocks in this economy is given by:

$$\frac{d^2 logY}{dlogA_j dlogA_i} = \frac{d\lambda_i}{dlogA_j} = \sum_k (\theta_k - 1)\lambda_k Cov_{\Omega_{(k)}}(\Psi_{(i)}, \Psi_{(j)}). \tag{8}$$

For a detailed derivation of Equation (8), see the appendix of Baqaee and Farhi (2019). To inspect how firm-level shocks can propagate through supply chains, consider the following example. Firm j, which is located in a *Red* zone, suffers a negative productivity shock, given by $d\log A_j < 0$. The second order term captures the reallocation effect: In response to a negative shock to product category j, all products k downstream of j may readjust their demand for all other inputs. Crucially, the impact of such readjustments by any given k on the output of product i depends on the size of product k as captured by its *Domar* weight λ_k, the elasticity of substitution θ_k in $k's$ production function, and the extent to which the supply chains that connect i and j to k coincide with one another, as given by the covariance term.

4.1 *Equations to estimate firm-level elasticity of substitution across suppliers*

The model yields estimating equations we use to estimate firm-level elasticities of substitution across suppliers within a product. We introduce a notation change to facilitate the exposition: a firm n can be either a buyer $b \in \mathcal{F}$ or a seller $s \in \mathcal{F}$. A firm b that sells product $j \in \mathcal{I}$ maximizes profits subject to its technology and to a CES bundle of intermediate inputs:

$$\max_{\{l_{bj}, x_{si,bj}\}} p_{bj} y_{bj} - w_{bj} l_{bj} - \sum_i \sum_s p_{si,bj} x_{si,bj}$$

subject to Equations (3), (4), and (5). Details about the optimization problem are in Appendix D.1. The maximization problem yields the following expression:

$$\log\left(\frac{PM_{si,bj}}{PM_{i,bj}}\right) = (1-\epsilon)\log\left(\frac{p_{si,bj}}{p_{i,bj}}\right) + \log\left(\mu_{si,bj}\right), \tag{9}$$

where $p_{i,bj} = \left(\sum_{s'}\left(p_{s'i,bj}^{1-\epsilon}\mu_{s'i,bj}\right)\right)^{\frac{1}{1-\epsilon}}$ is a CES price index, $PM_{si,bj} \equiv p_{si,bj} x_{si,bj}$, and $PM_{i,bj} \equiv \sum_s PM_{si,bj}$. $\log\left(\mu_{si,bj}\right)$ is the error term. This is the underlying basis for our estimation of the firm-level elasticity of substitution parameter ϵ, as will be described in detail in Section 5. Note that the results of this estimation procedure hold with any CES production function with an arbitrary number of nests as long as the lowest nest consists of suppliers within the same HS-4 product.

4.2 *Equations to estimate firm-level elasticity of substitution across products*

In this section, we derive conditions from the model to estimate the firm-level elasticity of substitution across products, as in some previous work (Atalay, 2017; Peter and Ruane, 2022; Boehm et al., 2019; Carvalho et al., 2021). We rewrite the maximization problem of the firm such that it maximizes

$$\max_{\{l_{bj}, x_{i,bj}\}} p_{bj} y_{bj} - w_{bj} l_{bj} - \sum_i p_{i,bj} x_{i,bj}$$

subject to Equations (3) , (4), and $p_{i,bj} = \left(\sum_s \mu_{si,bj} p_{si,bj}^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}$. Details on the optimization problem are in Appendix D.2.1. The maximization problem yields the following expression:

$$\log\left(\frac{PM_{i,bj}}{PM_{bj}}\right) = (1-\zeta)\log\left(\frac{p_{i,bj}}{p_{bj}}\right) + \log\left(w_{i,bj}\right), \tag{10}$$

where $p_{bj} = \left(\sum_{i'}\left(p_{i',bj}^{1-\zeta} w_{i',bj}\right)\right)^{\frac{1}{1-\zeta}}$ is a CES price index, $PM_{i,bj} \equiv p_{i,bj}x_{i,bj}$, and $PM_{bj} \equiv \sum_i PM_{i,bj}$. $\log\left(w_{i,bj}\right)$ is the error term. This is our estimating equation for the firm-level elasticity of substitution ζ which we take to the data, as described in Section 5.

5 Estimation Strategy

This section discusses how we estimate the primary elasticities in our model. The vector of parameters is $\widehat{\theta} = (\sigma, \alpha, \zeta, \epsilon)$. We set the elasticity of substitution between different consumption varieties $\sigma = 4$ (Broda and Weinstein, 2006), and the elasticity of substitution between labor and the composite intermediate input $\alpha = 0.5$ (Baqaee and Farhi, 2019). We now estimate the firm-level elasticity of substitution across suppliers (ϵ) and the firm-level elasticity of substitution across products (ζ) leveraging variation in the lockdown zones.

5.1 *Estimating equations for ϵ and ζ*

To estimate ϵ from Equation (9), the first major challenge we face is that the price index $p_{i,bj}$ includes the unobserved quantity $\mu_{si,bj}$ which denotes the importance of input from supplier s of product i in the production of buyer b producing product j. This unobserved quantity could depend on a number of factors such as unobserved input demand shocks or the buyer's preference for certain inputs. In order to construct changes in price indices that are observable, we follow Redding and Weinstein (2020) and assume that the overall importance of a product in a buyer's input use does not change between two consecutive months, even though the importance of inputs from suppliers within a product category can change.[18] We arrive at the equation below that links the overall expenditure share on a certain supplier s (as a share of total expenditure on product i) to the corresponding relative prices:

$$\log\left(\frac{\widehat{PM}_{si,bj,t}}{\widehat{PM}_{i,bj,t}}\right) = \omega_{b,t} + \omega_{i,t} + \omega_{b,i} + \omega_{s,i} + (1-\epsilon)\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right) + \log\left(\widehat{\lambda}_{i,bj,t}\widehat{\widetilde{s}}^{*}_{i,bj,t}\right) + \mathbf{X}\boldsymbol{\delta} + \xi_{si,bj,t}\ , \quad (11)$$

where $\widehat{v}_t = \frac{v_t}{v_{t-1}}$ are variables in changes with respect to the previous month. $\left[\omega_{b,t}, \omega_{i,t}, \omega_{b,i}, \omega_{s,i}\right]$ is a set of fixed effects, including buyer-by-month, product-by-month, buyer-by-product, and seller-

[18]This assumption simply requires that, for instance, a shoemaker's overall preference for leather in shoe manufacturing does not change, although its preference for leather from certain suppliers can change. Demand shocks may change $\mu_{si,bj,t}$ (e.g., the demand for leather from certain suppliers), but the geometric mean of $\mu_{si,bj,t}$ across suppliers within a product is stable between t and $t-1$. This enables us to construct changes in price indices that are not dependent on $\mu_{si,bj,t}$, but are directly observed in the data (details in Appendix D.1.2).

by-product fixed effects. $\widetilde{p}_{i,bj,t} = \prod_{s\in\Omega^*_{i,bj,t}} p_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}}$ is a geometric mean of unit values across common suppliers, where $\Omega^*_{i,bj,t} \equiv \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$ is the set of common suppliers for buyer b that appear in both the current and previous month, and $N^*_{i,bj,t} \equiv \Omega^*_{i,bj,t}$ is the number of common suppliers for buyer b in month t. Our setup has the advantage that we can decompose the change in price buyer b pays for inputs from seller s between $\widetilde{p}_{i,bj,t}$, the change in expenditure share $\widehat{\widetilde{s}}^*_{i,bj,t}$ and a Feenstra (1994) correction term $\widehat{\lambda}_{i,bj,t}$ that controls for the fact that sellers enter and exit in the data. More details are in Appendix D.1.3. Standard errors are two-way clustered at the origin and destination state levels. $\mathbf{X}$ are controls, including exposure to foreign demand and supply shocks, the number and severity of Covid cases, and geographic and cultural distance.

The wide array of high-dimensional fixed effects helps control for demand shocks (buyer-by-month fixed effects), product-level changes in demand or supply (product-by-month fixed effects), and buyer-by-seller and product-specific time-invariant characteristics (buyer-by-product and seller-by-product fixed effects). The remaining variation likely isolates time-varying changes across sellers within a product category. Yet, as we explain in the next section, we strengthen this framework by leveraging the mosaic of Covid-19 lockdowns to derive exogenous policy-induced variation in relative prices.

Now, to estimate ζ from Equation (10), there are two issues to address. First, notice that the price index $p_{i,bj,t}$ is a function of (unobservable) demand shocks $\mu_{si,bj,t}$, and ϵ, such that $p_{i,bj,t} \equiv \left(\sum_s \mu_{si,bj,t} p_{si,bj,t}^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}$. Second, the price index $p_{bj,t}$ is also a function of unobservable product-level demand shocks $w_{i,bj,t}$, which makes their computation challenging.

First, we construct price indices as $p_{i,bj,t} \equiv \left(\sum_s \mu_{si,bj,t} p_{si,bj,t}^{1-\widehat{\epsilon}}\right)^{\frac{1}{1-\widehat{\epsilon}}}$, where $\widehat{\epsilon}$ are estimated previously, $p_{si,bj,t}$ come directly from the data, and demand shocks $\mu_{si,bj,t}$ are constructed recursively. This recursive construction of demand shocks comes from predicting residuals from Equation (11) and setting an initial value for shocks $\mu_{si,bj,0}$ (Appendix D.2.2).

Second, we construct buyer-level price indices $p_{bj,t}$ following Redding and Weinstein (2020). We assume that the overall importance of the composite intermediates at the HS-4 level in the production function does not change between consecutive months. As such, we can construct this price independent of product-level demand shocks $w_{i,bj,t}$ after controlling for buyers' expenditure shares by product. More details about this are in Appendix D.2.1.

We then derive the following expression we take directly to the data:

$$\log\left(\frac{\widehat{PM}_{i,bj,t}}{\widehat{PM}_{bj,t}}\right) = \omega_{b,t} + \omega_{i,t} + \omega_{b,i} + (1-\zeta)\log\left(\frac{\widehat{p}_{i,bj,t}}{\widehat{\widehat{p}}_{bj,t}}\right) + \log\left(\widetilde{s}_{bj,t}\right) + \mathbf{X}\boldsymbol{\delta} + \xi_{i,bj,t} \ , \tag{12}$$

where $\left[\omega_{b,t}, \omega_{i,t}, \omega_{b,i}\right]$ are a set of buyer-by-month, product-by-month, and buyer-by-product fixed effects, which again account for a wide array of demand shocks, product shocks, and buyer-product specific characteristics. $\mathbf{X}$ are the same set of controls used before. $\widetilde{p}_{bj,t} \equiv \prod_{i=1}^{N_{bj,t}} \widetilde{p}_{i,bj,t}^{\frac{1}{N_{bj,t}}}$ is the geometric mean of unit values across products that buyer b purchases, and $\widetilde{s}_{bj,t} \equiv \prod_{i=1}^{N_{bj,t}} \widetilde{s}_{i,bj,t}^{\frac{1}{N_{bj,t}}}$ is the geometric mean of expenditure shares across products. Detailed derivations are in Appendix D.2.

5.2 *Addressing endogeneity concerns*

Despite the wide range of fixed effects, OLS estimates of ϵ may still be biased if additional unobserved demand-side shocks (changing $\mu_{si,bj,t}$) drive changes in prices and expenditure shares. The firm-level elasticity of substitution is a function of the slope of the buyer's input demand curve, and hence simultaneous shifts in the demand and supply curves induced by the Covid-19 shock would bias our estimates. For example, if Covid-19 induced demand shocks led to contractions in buyers' income and, at the same time, supply shocks led to contractions in the sellers' supply, the demand curves will look flatter (estimated ϵ higher) compared to the unbiased value of ϵ. Additionally, measurement error in input prices, proxied by unit values, may induce attenuation biases.

Our estimation strategy, therefore, involves using the sudden demarcations of lockdown zones that restrict economic activity in certain Indian districts as an instrumental variable when estimating this equation in two-stage least squares (2SLS). We use the disruptions in prices caused by sudden lockdowns that made it costlier for sellers in *Red* and *Orange* zones to produce and send their intermediate goods. The idea is that, after controlling for the wide array of fixed effects, the lockdown zones the buyer is located in, exposure to international demand and supply shocks, and the number and severity of regional Covid-19 cases, the remaining variation in prices facing a buyer are driven by supply shocks induced by policy-mandated sudden changes in the seller's lockdown zones. In addition, since the goods from the seller to the buyer have to transit through several districts located in different lockdown zones facing different severity in the movements of trucks and border controls, changes in the costs of transportation induced by these lockdowns provide another source of exogenous variation to estimate the firm-level elasticity of substitution.

To formalize the intuition behind our identification strategy, we assume that prices can be separated between prices at the origin and a trade cost, such that

$$\log\left(\widehat{p}_{si,bj,t}\right) = \log\left(\widehat{\tau}_{s,b,t}\right) + \log\left(\widehat{p}_{si,t}\right).$$

Here we can see the type of variation driving the two types of instruments we use. First, exogenous shifters to prices at the seller level $p_{si,t}$, such as economic restrictions induced by the lockdown zone the seller is located in, help us obtain unbiased estimates of the elasticity ϵ. Second, exogenous shifters at the seller-buyer level, for example, changes in transportation costs $\tau_{s,b,t}$ driven by the lockdown zones of the districts the goods pass through, also induce the needed variation. We now describe each of these instruments and then implement them within our estimation strategy.

Seller-level instruments. We derive supply-side shifters to obtain unbiased elasticities of substitution. Shocks induced by the Covid-19 lockdown policies that only impact sellers provide this variation. In Equation (13) below, we formalize this intuition.

$$log(\widehat{p}_{si,t}) = \beta^{R,p} Red_{o_{(s)}} Lock_t + \beta^{O,p} Orange_{o_{(s)}} Lock_t + \nu^{p}_{si,t}\ , \tag{13}$$

where $Lock_t$ is an indicator variable that equals 1 for the months from March to May of 2020, which are the months when the lockdown policies were implemented, 0 otherwise, and $Red_{o_{(s)}}$ and $Orange_{o_{(s)}}$ are indicator variables that equal 1 whenever seller s was located in *Red* or *Orange* districts, respectively.

Seller-Buyer-level instruments. The transportation of supplies from the location of the supplier to the buyer requires going through different districts, each of which is affected by lockdown policies in different ways. Intuitively, a route that contains more *Red* districts should increase the cost of transportation in contrast with a route with no *Red* districts. We construct instruments that capture that idea. We allow trade costs to change over time such that we can leverage the Covid-19 lockdown policy. In particular, as we describe in Appendix D.1.4, we assume

$$\log(\widehat{\tau}_{sb,t}) = \sigma \log(\widehat{TravelTime}_{sb,t}).$$

We leverage the Covid-19 lockdown as an exogenous shifter that only influences travel time between locations of seller s and buyer b, as reflected in Equation (14) below.

$$log(\widehat{\tau}_{sb,t}) = \beta^{R,\tau} Red_{o_{(s)}d_{(b)}} Lock_t + \beta^{O,\tau} Orange_{o_{(s)}d_{(b)}} Lock_t + \nu^{\tau}_{sb,t}. \tag{14}$$

Detailed derivations and estimating equations are in Appendix D.1.4. $Red_{o_{(s)}d_{(b)}}$ and $Orange_{o_{(s)}d_{(b)}}$ are the share of districts designated as *Red* and *Orange*, respectively, along the route between seller s and buyer b. We construct these variables using the Dijkstra algorithm for least-cost routes. Details about the implementation of this algorithm are in Appendix C.

Finally, we instrument the changes in relative prices in Equation (12) to estimate ζ. Potential unobservable product-level demand shocks could again induce an upward bias to OLS estimates of ζ. To construct our instruments, we leverage the seller-level and seller-buyer-level instruments we used to estimate ϵ and calculate weighted averages across suppliers to instrument for the change in relative prices for buyers. The intuition is that buyers that purchased inputs either from a larger share of sellers in *Red* zones, or from sellers located in districts such that the trading routes comprised of a large number of *Red* zone districts were more exposed to supply disruptions induced by the Covid-19 lockdowns. More details of the instruments are in Appendix D.2.3.

Discussion of instruments. The instruments induce buyers of certain types to be more affected than others based on their production networks. The Local Average Treatment Effect (LATE) may not represent the Average Treatment Effect (ATE) if buyers in *Red*, *Orange*, and *Green* zones already traded intensively with sellers in certain lockdown zones, and there is heterogeneity in responses. For instance, if buyers in *Red* traded mostly with sellers in *Red*, then our instrument may estimate effects on firms induced by having more *Red* sellers, and so it would upweight effects on buyers in *Red*. In Figure A2, we run two sets of distributional checks to investigate these patterns. These figures show that, in general, sellers from *Red*, *Orange*, and *Green* zones had similar interactions with buyers from *Red*, *Orange*, and *Green* zones.

We also consider whether certain products are sourced intensively from firms located in certain zones. For instance, if all the rubber supply of firms in this production network comes from suppliers in *Red* zones, then buyers of rubber would find it increasingly difficult to find suppliers. Once again, if there is heterogeneity in responses by product category, our estimated LATE elasticity would weigh rubber products higher than non-rubber products. While not a source of bias, it does affect the interpretation of the estimated parameter. In Figures A2g and A2h, we plot the shares of total purchases of each industry (HS Section) that are sourced from firms in *Red*, *Orange*, and *Green* zones. With the exception of the small HS industry 19 (arms and ammunition), there is no noticeable degree of concentration of suppliers from any particular zone.

6 ELASTICITY ESTIMATES

In this section, we show the results of the estimation of both firm-level elasticities of substitution across suppliers within a product category, and then across product categories.

6.1 *Firm-level elasticities of substitution across suppliers*

First, we report OLS estimates in Table 1. The implied elasticities exhibit values between 0.75 and 0.78 across the different specifications. In column (1), we include both buyer-by-month and product-by-month fixed effects. In column (2), we also include buyer-by-product and seller-by-product fixed effects. We obtain a similar elasticity of 0.78. To test whether our estimates vary by product aggregation, in columns (3) and (4), the estimates are based on 6-digit and 8-digit HS codes. The elasticities are around 0.75, so the estimates do not meaningfully change. Since these elasticities are below 1, these estimates suggest that, at the firm level, suppliers act as complements rather than substitutes for buyers. This is important for aggregate incomes since, from Equation (8) we can see that, once we consider second-order effects, an elasticity of substitution less than 1 implies that the aggregate impacts of negative shocks are amplified.

TABLE 1: OLS, firm-level elasticity of substitution across suppliers

	(1)	(2)	(3)	(4)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.2171	0.2222	0.2506	0.2441
	(0.0133)	(0.0147)	(0.0324)	(0.0352)
ϵ	0.7828	0.7777	0.7493	0.7558
R^2	0.4177	0.4601	0.4838	0.4958
Obs	2028039	1966591	851483	993583
HS digits	4	4	6	8
Buyer-month FE	Y	Y	Y	Y
Product-month FE	Y	Y	Y	Y
Buyer-product FE		Y	Y	Y
Seller-product FE		Y	Y	Y

Notes: OLS estimates come from estimating Equation (11). The first row reports the estimates associated with changes in log relative unit values. Standard errors are two-way clustered at the origin and destination state level and are reported in parentheses below each estimate. The third row reports the implied value for ϵ, which is 1 minus the estimate on the first row. The table contains four columns: Each corresponds to different specifications on how we define a product (4-digit, 6-digit, or 8-digit HS codes) and which fixed effects we include. These combinations are reported by the last five rows of the table. All specifications include the controls mentioned in the paper.

Nevertheless, as we describe previously, it is possible that the OLS estimates are contaminated by simultaneous demand shocks that happened during Covid-19. In Table 2, we report 2SLS estimates based on our proposed instruments. We find evidence that inputs across different suppliers of a firm within the same 4-digit HS product category are highly complementary, ranging from $0.50-0.66$, depending on the set of fixed effects and instruments we use. In Table A4, we show similar patterns for HS-6 and HS-8 categories. Our preferred specification is column (3) with an elasticity of 0.55, where we use both the seller and the seller-buyer level instrument, essentially deriving variation from both sellers' production costs and transportation costs. We include buyer-by-month and product-by-month fixed effects that account for time-varying demand shocks, and also account for firm entry and exit with the Feenstra (1994) term. Each specification reports a high Kleibergen-Paap F-statistic, indicating that our instruments are statistically relevant. In columns (1) and (2), we use the seller-level and seller-buyer-level instruments separately. The elasticities are 0.50 and 0.61, respectively, which also reflect complementarity. Finally, in column (4), we also include buyer-by-product and seller-by-product fixed effects, and the elasticity rises to 0.66.

TABLE 2: 2SLS, firm-level elasticity of substitution across suppliers

	(1)	(2)	(3)	(4)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.5042	0.3945	0.4538	0.3409
	(0.2129)	(0.0933)	(0.1389)	(0.1068)
ϵ	0.4957	0.6054	0.5461	0.6590
Obs	2854292	2028039	2028039	1966591
K-PF	48.232	133.688	143.413	248.977
Seller IV	Y		Y	Y
Bilateral IV		Y	Y	Y
Buyer-month FE	Y	Y	Y	Y
Product-month FE	Y	Y	Y	Y
Buyer-product FE				Y
Seller-product FE				Y

Notes: 2SLS estimates come from estimating Equation (11). The set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$. That is, a supplier s of buyer b is considered *common* if they traded in both the current and previous month. The first stage uses either bilateral or seller-level instruments, as reported in rows six and seven. Bilateral instruments correspond to Equation (14), and seller-level instruments correspond to Equation (13). The first row reports estimates associated with changes in relative unit values in logs. Standard errors are two-way clustered at the origin and destination state level and are reported in parentheses below each estimate. The third row reports the implied value for ϵ, which is 1 minus the estimate on the first row. The fourth row reports the number of observations. The fifth row reports the Kleibergen-Paap F statistic from the first stage. A product is a 4-digit HS code, and the treatment period is March-May 2020. The table contains four columns. Each column corresponds to different combinations of instruments and fixed effects. These combinations are reported in the last six rows. All specifications include the controls mentioned in the paper.

The 2SLS estimates for ϵ are smaller than the OLS estimates. As discussed in Section 5.2, the bias is in the expected direction if we expect the Covid-19 shock to also induce negative demand shocks, thereby biasing up OLS estimates of ϵ. We may expect that our estimated elasticity will be lower for the sub-sample of buyers who did not have more than one supplier to source inputs from. In Table A3, we restrict our sample to cases when a buyer traded with at-least two sellers in two consecutive periods. Column (3), our preferred specification, yields an elasticity of substitution of 0.59, very close to the estimate from our main specification.

Elasticities by Product Aggregation. To examine differences by the level of aggregation of the product, we re-estimate our main specification in Table A4 using HS-6 and HS-8 as product definitions. Finer product classifications (e.g., HS-8) may imply that there are fewer suppliers one may be able to source from, and so we may expect a lower elasticity of substitution between suppliers. In columns (1) and (3) we replicate our main specifications, with elasticities of 0.43 (for HS-6) and 0.06 (for HS-8), respectively. These numbers reflect even higher degrees of complementarity when we consider a more granular notion of product. Overall, these patterns suggest that inputs are highly specific for buying firms.

Elasticity Heterogeneity by Industry. We analyze whether the degree of substitution across suppliers varies by industry (HS section). Firms that source from highly specific intermediate inputs (i.e. processed foods) should report a lower elasticity of substitution across suppliers than firms that source from more general inputs (e.g. textiles). In Table A5 and Figure A4, we show the estimates of this elasticity of substitution across suppliers by twenty-one broad industries. We find that the OLS elasticity of substitution across suppliers by industry lies in the range of $0.41-0.87$. Once we instrument for the unit values with the Covid-19-induced lockdown variation, we find that there is wider heterogeneity across industries. Indeed, we find that that *Processed foods* yield an elasticity of 0.19, while *Textiles* yield an elasticity of 0.81. Also, while for the majority of the industries we find evidence of complementarity, there are some industries such as Plastics, Vegetables, and Handicrafts where suppliers within an HS-4 product are likely substitutes.

Elasticity heterogeneity by Institutional Quality. After estimating our firm-level elasticities of substitution across suppliers, a natural question is to inquire about a key mechanism behind our results. We argue that complementarity in production at the firm level arises because of the nature of the production of intermediate inputs. The production of these inputs ranges from homogeneous (i.e., commodities) to relationship-specific (Rauch, 1999). On one side, homogeneous goods are expected to be close substitutes. Conversely, relationship-specific goods are expected to exhibit complementarity at the firm level. Since relationship-specific inputs require more highly specified contracts for their production, they are more likely to be traded in locations with better institutional quality (e.g., contract enforcement). Therefore, we may expect to find our complementarity result to hold for inputs produced or purchased by firms in locations with better institutional quality (Boehm, 2022; Boehm and Oberfield, 2020).

To test our hypothesis, we use the geographic variation in institutional quality across Indian districts. From the *Socioeconomic High-resolution Rural-Urban Geographic Platform for India* (SHRUG) database, we calculate the average time it takes for courts in each Indian district to reach a verdict. In particular, we measure low (high) institutional quality as districts that take longer (shorter) time to reach a verdict. We classify districts into terciles (*Low*, *Medium*, and *High*), and we estimate our firm-level elasticities of substitution across suppliers conditional on whether sellers or buyers are located in each of these terciles of institutional quality.

In Table A6, we report 2SLS estimates of our elasticity conditional on the institutional quality. The first three columns focus on sellers, and the last three on buyers. If we focus on the columns for sellers, as expected, we find an elasticity of substitution across suppliers located in low-quality districts of $1.59 > 1$ (substitution), and an elasticity for suppliers located in high-quality districts of $0.23 < 1$ (complementarity). The substitution across suppliers located in medium-quality districts is 0.73, which is just in between the previous two estimates. We find a similar picture if we focus on buyers, where we find mild substitutability for buyers located in low-quality districts ($1.04 > 1$), and complementarity for buyers located in high-quality districts ($0.72 < 1$).

Longer-run elasticities of substitution across suppliers. So far, we provide evidence that the estimates of the firm-level elasticity of substitution across suppliers exhibit complementarity in production ($\epsilon = 0.55 < 1$). More precisely, we estimate a short-run elasticity since the frequency of the data is monthly. That is, we allow firms a month to adjust expenditures in response to price shocks. A potential avenue of exploration is to understand how this elasticity changes with the frequency of the measurement. As such, we estimate our elasticities after aggregating the data at the quarterly level; allowing firms a quarter to adjust expenditures in response to price shocks.

In Table A7, we report OLS and 2SLS estimates of the firm-level elasticity of substitution across suppliers at both frequencies. The first two columns show our baseline estimates of ϵ at a monthly frequency. They correspond to column 1 from Table 1 and column 3 from Table 2, respectively. The last two columns show the estimates of ϵ at a quarterly frequency. As expected, the estimates at quarterly frequency are slightly higher in comparison to the monthly frequency ones: the OLS estimates go up from 0.78 to 0.86, and the 2SLS estimates go up from 0.55 to 0.79. More importantly, our quarterly estimates still exhibit complementarity in production since they are significantly below 1.

6.2 *Firm-level elasticities of substitution across products*

In Table 3, we report our estimates for the firm-level elasticity of substitution across products. In column (1), we show the OLS estimate of $\zeta = 0.92$, which reflects complementarity between product categories. In columns (2) and (3), we define products more granularly. In this case, the elasticities are around 0.80, which also reflects complementarity between products.

TABLE 3: Firm-level elasticity of substitution across products

	(1)	(2)	(3)	(4)	(5)	(6)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.0842	0.2014	0.1996	0.3136	0.1712	0.1996
	(0.0039)	(0.0045)	(0.0048)	(0.1060)	(0.0040)	(0.0048)
ζ	0.9157	0.7985	0.8003	0.6863	0.4368	0.4721
Obs	1292329	794376	766804	1292329	794376	766804
K-PF	.	.	.	27.284	17.950	15.868
Estimator	OLS	OLS	OLS	2SLS	2SLS	2SLS
HS digits	4	8	8	4	8	8
Product-month FE	Y	Y	Y	Y	Y	Y
Buyer-month FE	Y	Y	Y	Y	Y	Y
Buyer-product FE			Y			Y

Notes: The estimates come from estimating Equation (10). Price indices are constructed by recovering the residuals when estimating ϵ and the estimates of ϵ. The first three columns are OLS estimates of ζ. The last three columns are 2SLS of estimates of ζ using weighted averages of both bilateral and seller-level instruments across sellers. Bilateral instruments correspond to Equation (14), and seller-level instruments correspond to Equation (13). Each column corresponds to a different combination of fixed effects and HS codes. Columns (1)-(2) and (4)-(5) correspond to our preferred specification when estimating ϵ for 4-digit and 8-digit HS codes. Additionally, in columns (3) and (6), we also include buyer-by-product fixed effects. The first row reports the estimates associated with changes in log relative unit values. Standard errors are clustered at the buyer's district level and are reported in parentheses below each estimate. The third row reports the implied value for ζ, which is 1 minus the estimate on the first row. The fourth row reports the number of observations. The fifth row reports the Kleibergen-Paap F statistic from the first stage. The sixth row denotes whether the estimates are obtained through OLS or 2SLS. The seventh row reports the HS code. The last three rows indicate the combination of fixed effects.

In columns (4)-(6) we report our estimates of ζ under 2SLS estimation after using a weighted average of instruments across buyers' sellers as discussed in Section 5.2. Our specification in column (4) reports a value of 0.69, reflecting that simultaneous negative demand and supply shocks during Covid-19 led to an underestimation of ζ under OLS. This elasticity is higher than the 2SLS elasticity of substitution across suppliers for the same product ($\epsilon = 0.55$), reflecting a lower degree of complementarity across products compared to suppliers.[19] In columns (5) and (6), similar values for this elasticity hold when we define a product as 8-digit HS codes, and after the inclusion of buyer-by-product fixed effects. Finally, first-stage F-stats are high, which reflects the statistical relevance of our weighted averaged instruments.

Unlike the elasticity of substitution across suppliers within a product category, there have been previous attempts in the literature to estimate the elasticity of substitution across products or

[19]This is consistent with the macroeconomics literature (Houthakker, 1955; Bachmann et al., 2022; Lagos, 2006).

industries. In particular, other work has estimated a wide range of values for parameters akin to ζ depending on the aggregation of the industry and on the research question. Our elasticity is close to Boehm et al. (2019) who estimate an elasticity across HS-10 products that lies between $0.42-0.62$ for non-Japanese affiliates and 0.20 for Japanese affiliates. Atalay (2017) finds an estimate of around 0.10 for 30 aggregated industries using US data.

7 Quantification and counterfactuals

In this section, we use both data from our production network and our newly estimated elasticities to quantify the role of these elasticities in the propagation of shocks. To do this, we need to write down the Leontief matrix in standard form. Given the production structure of our economy, we need four submatrices: (i) firm purchases of 4-digit HS products, (ii) firm sales of 4-digit HS products, (iii) labor employed by each firm, and (iv) final sales by each firm. The first two submatrices are directly constructed from the firm-to-firm trade data from the pre-Covid period of March 2019 to February 2020. Labor employed and final sales by firms are obtained by merging in firm-level data from Indiamart, which contains information on firm-level employment and final sales.[20] For more details for this, see Appendix C.

The economy is comprised of $N = 93,260$ firms, $I = 1,293$ different HS-4 products, labor, and a composite final good. The average firm buys 10 distinct products as a buyer, and sells 5 distinct products as a seller. The most connected buyer and seller buys and sells over 500 distinct products. We use this $94,555 \times 94,555$ input-output matrix to understand how complementarities at the firm level affect the propagation of shocks through production networks.

While recent work also quantifies the effect of firm-level shocks on aggregate GDP up to the second order, they mostly rely on changes in firm-level final sales rather than the full production network. Instead, we identify the most connected firms, deriving the Leontief inverse from the entire production network. Without relying on any approximation, we use the full network to quantify the importance of firm connectivity separately from firm size. This exponentially increases computational complexity from the order of $(N+I+2)$ to $(N+I+2)^2$. As such, we use computational innovations in big data to implement this procedure. For more details on the derivation of the shock propagation equation and its numerical implementation, see Appendix E.

Note that our quantification exercises in this section are conditional on the products that firms buy or sell being given at the extensive margin, even though a firm can change its set of buyers or suppliers (Khanna et al., 2022). We, therefore, need to empirically assess whether the set of

[20] https://www.indiamart.com/

HS-4 products a buyer buys and the set of HS-4 products that a seller sells, changes between the pre and the post-Covid period. We do this by inspecting whether both sellers and buyers of each product continued to trade in their corresponding product categories after the Covid-19 lockdowns. In Figure A5, we show the product-level distribution of the share of sellers that sold and buyers that purchased goods of that product during both time periods t and $t-1$, where t is a 6-month window before and after the lockdowns. In the figure, we see that, for both sellers and buyers, these two distributions are very similar to each other. The overall stability in Figure A5 shows that the assumption that the products that firms buy or sell do not change is tenable when analyzing the impact of negative productivity shocks.

7.1 *How important is the firm-level elasticity of substitution across suppliers?*

We assess the importance of the estimated firm-level elasticity of substitution across suppliers for the same product by studying how this elasticity determines the impacts of negative firm-level productivity shocks on aggregate GDP. In this counterfactual, we shock the productivity of firms located in *Red* zones by 25%. We find that this productivity shock reduces GDP by 10.95%. As an empirical benchmark, the state's annual GDP fell by 11.3% in 2020/21. This fall would be $2.68pp$ less in a model where firms in the same HS-4 product are considered substitutes ($\epsilon = 2$) and $0.99pp$ more when firms in the same HS-4 product are considered almost Leontief ($\epsilon = 0.001$). Given the quarterly GDP of this state in 2020-2021, the additional losses due to firm-level complementarities translate into 870 million USD, which is about 25 USD per capita per quarter, compared to the case when firms are substitutes.[21]

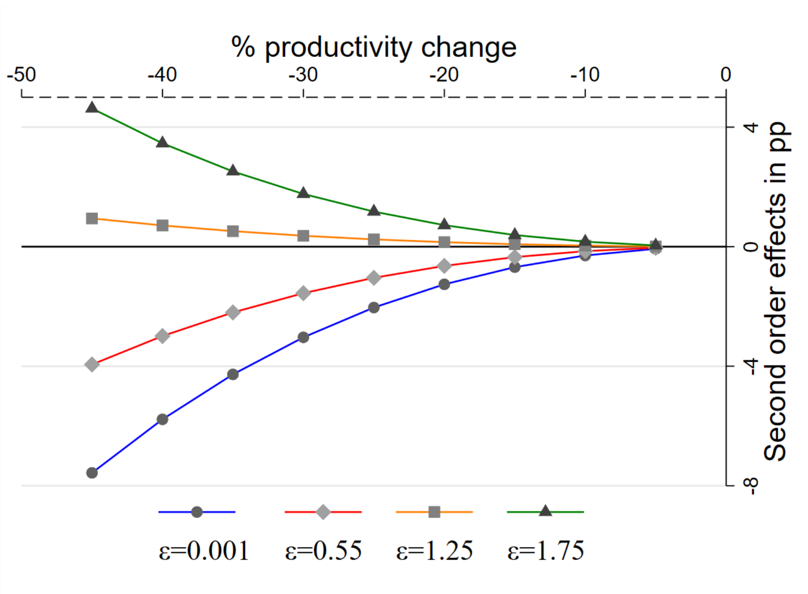
Note that the differences in GDP that arise from changing the values of firm-level elasticities of substitution across suppliers only change the second-order effects on GDP. Then, how important are these second-order effects that we have estimated? In Figure 5, we simulate different levels of negative productivity shocks for four different values of the elasticity ϵ and plot the second-order percentage point change in GDP due to these shocks. The blue and red lines show these differences for high levels of complementarity between suppliers: 0.001 and our estimated elasticity 0.55, respectively. The green and yellow lines show the additional second-order change in GDP for high levels of substitution across suppliers: 1.75 and 1.25, respectively.

Jointly, these plots provide two main lessons. First, for a given negative productivity shock, the second-order effects intensify with the degree of complementarity between suppliers. Second, given the same value of ϵ, the second-order effects intensify with the magnitude of the productiv-

[21] To put these numbers into perspective, Baqaee and Farhi (2019) showed that complementarities at the industry level, with an elasticity of substitution 0.001, amplify the effect of a negative 13% shock in the oil industry on GDP by around 0.61%.

ity shocks. Finally, as suppliers exhibit higher substitutability, the second-order effects actually dampen the negative first-order effects, and more so, for higher values of productivity shocks. When suppliers instead exhibit complementarity, the second-order effects magnify the negative first-order effects. That is, unlike the first-order effects, which only depend on firm size, complementarities at the firm level non-linearly amplify the effects of negative productivity shocks. This reflects similar amplification patterns that (Baqaee and Farhi, 2019) document, but at the industry level. These graphs illustrate the importance of second-order effects largely driven by firm complementarities, especially for large, short-lived, negative productivity shocks such as Covid-19.

FIGURE 5: How important are second-order effects?



Notes: The horizontal axis is the percentage change in productivity for firms in *Red* districts. The vertical axis is the second-order change in GDP in percentage points for different values of the firm-level elasticity of substitution across suppliers (ϵ). Different values of the elasticity ($\epsilon = 0.001$, $\epsilon = 0.55$, $\epsilon = 1.25$, and 1.75) are plotted with different colors.

7.2 *How important is a firm's connectivity in its network?*

Since (Hulten, 1978), policy-makers and researchers have emphasized the importance of firm sizes in the propagation of shocks. Now we investigate the importance of firm connectedness, given fixed firm sizes, in the propagation of shocks. We conduct these counterfactuals for different values of elasticities of substitution when suppliers are complementary, as our empirical analysis points to strong complementarities between suppliers.

In times of crises, governments often help small firms stay in business by providing subsidies.

Nevertheless, governments have limited funds, and it is often unclear how to best allocate their fixed budget among similarly sized firms. In this counterfactual, we explore the importance of a firm's connectivity in its network when implementing these subsidies. We measure the connectivity of a firm by its value within the Leontief inverse matrix, which measures firms' direct and indirect connections to other firms.[22] Firm size is measured by the size of its Domar weight.

Since firm sizes and connectivity are highly correlated with a correlation coefficient of 0.75, we vary the firms' connectivity for a given level of firm size to tease out the pure effect of connectivity. To implement this, we choose firms with Domar weights equal in size for up to 5 decimal places. The first set consists of the most connected firms, the second set is a random draw of firms, and the third set consists of the least connected firms. Since firm sizes are given, the first-order effects are the same irrespective of how connected the firms are.[23] In figure 6, we only plot the second-order effects on GDP under these three experiments. In the first scenario, only the most connected firms are affected by negative productivity shocks (blue line). In the second, a random draw of firms is affected (green line). Finally, in the third scenario, only the least connected firms are affected (red line). We perform these experiments under three different elasticities of substitution: an elasticity of substitution amounting to near perfect complementarity ($\epsilon = 0.001$) in the left panel, our estimated complementarity ($\epsilon = 0.55$) in the right panel, and near Cobb-Douglas ($\epsilon = 0.98$) in the bottom panel. All these experiments are conditional on given firm sizes; that is, we vary the connectivity of firms after matching on firm sizes.

These counterfactuals show that the fall in GDP is much larger if the most connected firms are affected compared to the least connected firms, or a random set of firms for a given firm size (Domar weight).[24] The importance of the most connected firms increases non-linearly with the negative productivity shocks: as the shock gets larger, it becomes increasingly important to give attention to the most connected firms. Our experiment suggests that for our baseline value of elasticity of substitution ($\epsilon = 0.55$) and a negative productivity shock of 45%, if governments save better-connected firms, given the same firm sizes, compared to randomly targeting firms, the fall in GDP would be $0.20pp$ less, and $0.31pp$ less compared to targeting the least connected firms.

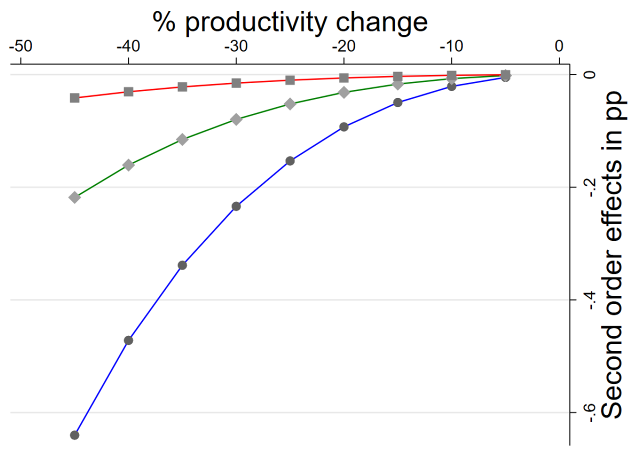
[22]The $(i, j)^{th}$ entry of the Leontief is a measure of firm i's total reliance on j as a supplier. Summing across all i's yields a measure of the connectivity of each supplier j or its importance in the firm network in terms of connectivity.

[23]This is the only counterfactual where we draw firms from the entire state rather than only firms located in *Red* zones to maximize the number of firms that vary in connectivity for a given firm size.

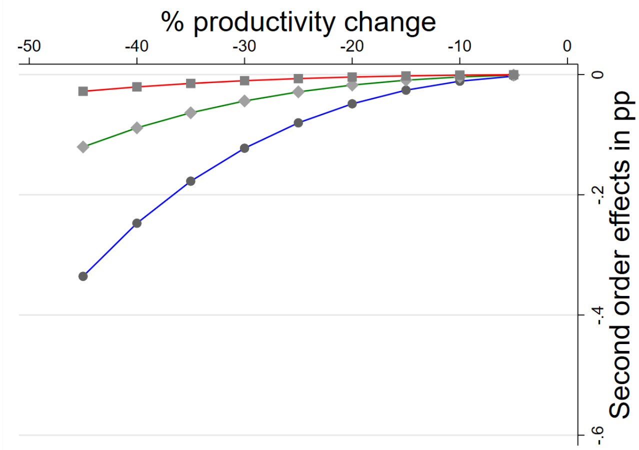
[24]Relatedly, Liu and Tsyvinski (2020) show that negative shocks to upstream sectors can have more adverse effects on GDP despite having identical Domar weights. While our effects materialize from the second-order propagation through a roundabout production network with firm complementarity, the results in Liu and Tsyvinski (2020) stem from adjustment costs in a vertical production network.

FIGURE 6: Second order GDP effects when firms with the same size but different connectivity are affected

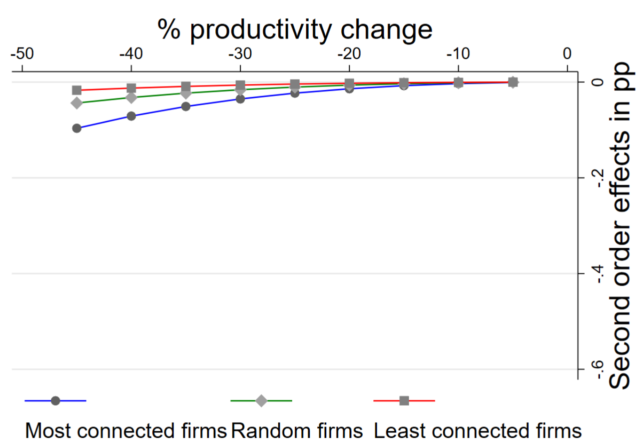
(a) $\epsilon = 0.001$



(b) $\epsilon = 0.55$



(c) $\epsilon = 0.98$



Notes: The figure comprises three panels. In each panel, the horizontal axis is the percentage change in productivity for firms in our state, and the vertical axis is the second-order change in GDP, in percentage points for different values of firm-level elasticity of substitution across suppliers (ϵ). Each panel corresponds to different values of the firm-level elasticity of substitution across suppliers (ϵ). Each panel contains three scenarios. The first scenario is the blue line, where the most connected firms are shocked. The second scenario is the green line, where random firms are shocked. The third scenario is the red line, where the least connected firms are shocked.

We notice three patterns: First, in the near Cobb-Douglas case ($\epsilon = 0.98$), the differences in GDP when bailing out the least or the most connected firms are negligible because the second-order effects are negligible.

Second, as the level of productivity shock increases, it becomes more important to save the most connected firms. While for a low productivity shock of 5%, the differences in GDP are negligible (0.001pp and 0.002pp), for a productivity shock of 25%, these differences are 0.05pp and 0.07pp compared to saving randomly connected and the least connected firms.

Third, the effects of these non-linearities are more pronounced when suppliers are highly complementary. For near-perfect complementarity ($\epsilon = 0.001$) and a high negative productivity shock (–45%), the gains from saving the most connected firms compared to saving randomly targeted and least connected firms are 0.38pp and 0.60pp, which is almost double the gains if instead suppliers were moderately complementary ($\epsilon = 0.55$).

If the policy's goal is to reduce the effects of negative productivity shocks on GDP, for large productivity shocks and low levels of elasticities of substitution, more effective subsidies should target highly connected firms, given firm sizes.

7.3 How important is measuring a firm's total (direct plus indirect) connectivity?

The existing literature has shown that shocks to a firm's suppliers affect the buyer firm and its suppliers (Barrot and Sauvagnat, 2018). There is also recent evidence that shocks to a firm can affect its direct as well as other indirect connections (Carvalho et al., 2021). In this counterfactual, we quantify how important it is to take into account a firm's indirect connectivity in understanding how shocks to the firm can affect aggregate GDP. To be precise, a firm's indirect connections measure not only the number of direct buyers of a supplier but also the buyers' buyers and their buyers and so on.[25]

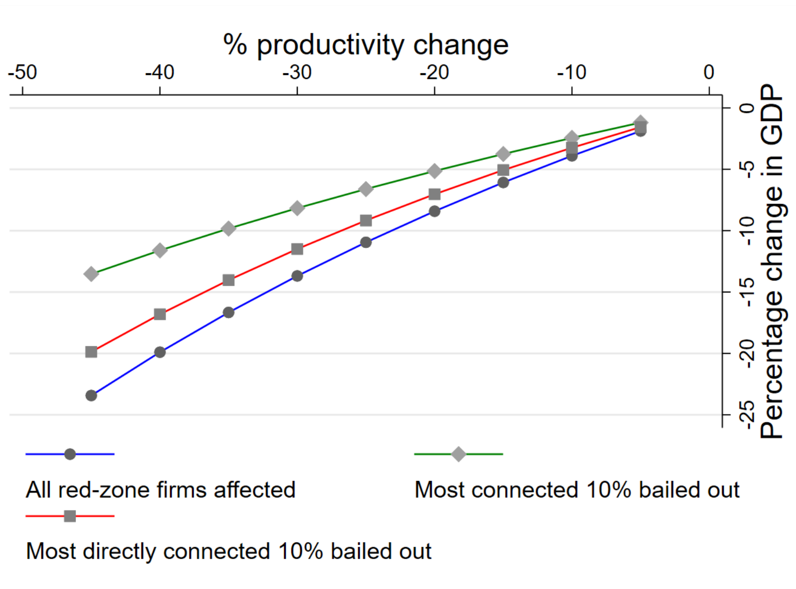
We conduct two experiments. In the first experiment, the government bails out the most directly connected 10% firms in *Red* zones, where direct connectivity is measured by the number of buyers a supplier directly supplies (red line in Figure 7). In the second experiment, the government bails the most connected 10% firms in *Red* zones, where the total connectivity of a firm is measured by all its direct and indirect connections (green line in Figure 7). Note that, unlike the previous counterfactual, we do not fix firm sizes and vary total connectivity. We are interested in understanding if the government were to bail out just the most directly connected firms as opposed to bailing out the most connected firms irrespective of size, how would that affect aggregate GDP. We report the total effect on GDP under these two sets of experiments and the baseline results (shock to all firms in *Red* zones).

We find that, under our estimated elasticity of $\epsilon = 0.55$ and a negative productivity shock of 25%, the fall in GDP would be 2.56pp less if the government were to pick firms on the basis of total

[25] As a reminder, we measure the total connectivity of a firm by its value within the Leontief inverse matrix, which measures firms' direct and indirect connections to other firms.

connectivity as opposed to direct connectivity. We see that as the level of the negative productivity shock increases, the difference in aggregate GDP between these two sets of experiments rises, emphasizing the importance of measuring a firm's indirect connections as well.

FIGURE 7: How important is a firm's connectivity in its network?



Notes: The horizontal axis is the percentage change in productivity for firms in *Red* districts. The vertical axis is the second-order change in GDP in percentage points for $\epsilon = 0.55$. The blue line corresponds to the baseline case when all firms in the *Red* districts are affected. The red line corresponds to the case when the government only bails out the 10% most directly connected firms. The green line corresponds to the case when the government bails out the 10% firms with the most (direct+indirect) total connections.

8 CONCLUSION

In this paper, we use highly disaggregated firm-to-firm transaction data from a large Indian state and provide one of the first estimates of elasticities of substitution across suppliers within the same product category at the firm level. We provide new estimation strategies and estimates for these elasticities by leveraging regional variation in supply-side shocks induced by the Indian government's massive lockdown policy. We find that suppliers of inputs are highly complementary even at this very granular level. This elasticity crucially determines aggregate impacts and the transmission of shocks across the network, but has previously eluded the literature (Baqaee and Farhi, 2019). The combined advantage of having product-level unit values and quasi-experimental variation in supply-side shocks allows us to overcome previous challenges in the literature, and credibly estimate this elasticity across suppliers of a particular product.

Since inputs are complementary, adverse shocks to even a small subset of firms that are highly linked in the supply chain can negatively affect the aggregate economy by propagating through firm networks. When we conservatively shock only the productivity of firms located in *Red* zones by 25%, we find that if suppliers of the same product were substitutes instead of complements, the fall in aggregate quarterly GDP in the state under study would be about 870 million USD lower, or about 25 USD per capita lower per quarter. Using big data computational techniques, we quantify this decline *directly* using information on the economy-wide firm-to-firm network without relying on first-order approximations. Our methods thus provide new techniques to quantify shocks through large and complex production networks. Using data on the entire production network in the state, we measure the full connectivity of firms in the network and show that as the level of complementarity and the magnitude of the negative productivity shock increase, it is more effective to save the more connected firms, after controlling for firm size.

Our findings have implications for policymakers worldwide, who often face difficult trade-offs in crisis regarding which firms to bail out. Given the underlying variation used, these estimates are relevant for other crises that are expected to remain short-lived, such as natural disasters, temporary trade wars and sanctions, and supply-chain disruptions.

References

Alfaro-Urena, A., I. Manelici, and J. P. Vasquez (2020). The Effects of joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages. *Available at SSRN 3376129*.

Alviarez, V., M. Fioretti, K. Kikkawa, and M. Morlacco (2021). Two-Sided Market Power in Firm-to-Firm Trade. *Working paper*.

Arkolakis, C., F. Huneeus, and Y. Miyauchi (2023). Spatial production networks. Technical report, National Bureau of Economic Research.

Atalay, E. (2017). How Important are Sectoral Shocks? *American Economic Journal: Macroeconomics 9*(4), 254–80.

Bachmann, R., D. Baqaee, C. Bayer, M. Kuhn, A. Löschel, B. Moll, A. Peichl, K. Pittel, M. Schularick, et al. (2022). What if? The Economic Effects for Germany of a Stop of Energy Imports from Russia. *ECONtribute Policy Brief 28*, 2022.

Baqaee, D. and E. Farhi (2020). Nonlinear Production Networks with an Application to the Covid-19 Crisis. *NBER Working Paper 27281*.

Baqaee, D. R. and E. Farhi (2019). The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem. *Econometrica 87*(4), 1155–1203.

Barrot, J.-N. and J. Sauvagnat (2018). Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks. *The Quarterly Journal of Economics 131*(3), 1543–1592.

Bas, M., A. Fernandes, and C. Paunov (2023). How Resilient Was Trade to COVID-19? *Economics Letters*, 111080.

Behrens, K., G. Corcos, and G. Mion (2013). Trade Crisis? What Trade Crisis? *Review of Economics and Statistics 95*(2), 702–709.

Beyer, R., T. Jain, and S. Sinha (2021). Lights out? Covid-19 Containment Policies and Economic Activity. *Covid-19 Containment Policies and Economic Activity*.

Boehm, C. E., A. Flaaen, and N. Pandalai-Nayar (2019). Input Linkages and the Transmission of Shocks: Firm-level Evidence from the 2011 Tōhoku Earthquake. *Review of Economics and Statistics 101*(1), 60–75.

Boehm, J. (2022). The Impact of Contract Enforcement Costs on Value Chains and Aggregate Productivity. *The Review of Economics and Statistics 104*(1), 34–50.

Boehm, J. and E. Oberfield (2020). Misallocation in the Market for Inputs: Enforcement and the Organization of Production. *The Quarterly Journal of Economics 135*(4), 2007–2058.

Bonadio, B., Z. Huo, A. A. Levchenko, and N. Pandalai-Nayar (2021). Global Supply Chains in the Pandemic. *Journal of International Economics*, 103534.

Bricongne, J.-C., L. Fontagné, G. Gaulier, D. Taglioni, and V. Vicard (2012). Firms and the Global Crisis: French Exports in the Turmoil. *Journal of International Economics 87*(1).

Broda, C. and D. E. Weinstein (2006). Globalization and the Gains from Variety. *The Quarterly Journal of Economics 121*(2), 541–585.

Cakmakli, C., S. Demiralp, S. Kalemli-Ozcan, S. Yesiltas, and M. Yildirim (2021). The Economic Case for Global Vaccinations: An Epidemiological Model with International Production Networks. *NBER Working Paper 28395*.

Carvalho, V. M., M. Nirei, Y. U. Saito, and A. Tahbaz-Salehi (2021). Supply Chain Disruptions: Evidence from the Great East Japan Earthquake. *The Quarterly Journal of Economics 136*(2).

Chacha, P., B. Kirui, and V. Wiedemann (2022). Mapping kenya's production network transaction by transaction. *Available at SSRN 4315810*.

Chakrabati, A., K. Mahajan, and S. Tomar (2021). Aggregate Shocks, Domestic Trade Collapse and Regional Realignment. *Available at SSRN 3869005*.

Demir, B., A. C. Fieler, D. Xu, and K. K. Yang (2021). O-ring Production Networks. *NBER Working Paper 28433*.

Demir, B. and B. Javorcik (2020). Trade Finance Matters: Evidence from the Covid-19 Crisis. *Oxford Review of Economic Policy 36*, S397—-S408.

Dew-Becker, I. (2022). Tail risk in production networks. Technical report, National Bureau of Economic Research.

Dhyne, E., A. K. Kikkawa, and G. Magerman (2022). Imperfect Competition in Firm-to-Firm Trade. *Journal of the European Economic Association 20*(5), 1933–1970.

Dhyne, E., A. K. Kikkawa, M. Mogstad, and F. Tintelnot (2021). Trade and Domestic Production Networks. *The Review of Economic Studies 88*(2), 643–668.

Edmond, C., V. Midrigan, and D. Y. Xu (2018). How Costly are Markups? *NBER Working Paper 24800*.

Fajgelbaum, P. D. and E. Schaal (2020). Optimal Transport Networks in Spatial Equilibrium. *Econometrica 88*(4), 1411–1452.

Feenstra, R. C. (1994). New Product Varieties and the Measurement of International Prices. *The American Economic Review*, 157–177.

Ferrari, A. (2022). Inventories, Demand Shocks Propagation and Amplification in Supply Chains. *arXiv preprint arXiv:2205.03862*.

Gerschel, E., A. Martinez, and I. Mejean (2020). Propagation of Shocks in Global Value Chains: The Coronavirus Case. *Notes IPP* (53).

Giovanni, J. d. and A. A. Levchenko (2009). Trade Openness and Volatility. *The Review of Economics and Statistics 91*(3), 558–585.

Heise, S. (2020). How did China's Covid-19 Shutdown affect US Supply Chains? *Liberty Street Economics - Federal Reserve Bank of New York*.

Houthakker, H. S. (1955). The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis. *The Review of Economic Studies 23*(1), 27–31.

Hulten, C. R. (1978). Growth Accounting with Intermediate Inputs. *The Review of Economic Studies 45*(3), 511–518.

Hummels, D., R. Jørgensen, J. Munch, and C. Xiang (2014). The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm data. *American Economic Review 104*(6).

Huneeus, F. (2018). Production network dynamics and the propagation of shocks. *Graduate thesis, Princeton University, Princeton, NJ*.

Jones, C. I. (2011). Intermediate Goods and Weak Links in the Theory of Economic Development. *American Economic Journal: Macroeconomics 3*(2), 1–28.

Khanna, G., N. Morales, and N. Pandalai-Nayar (2022). Supply Chain Resilience: Evidence from Indian Firms. *Working Paper*.

Kone, Z. L., M. Y. Liu, A. Mattoo, C. Ozden, and S. Sharma (2018). Internal Borders and Migration in India. *Journal of Economic Geography 18*(4), 729–759.

Korovkin, V. and A. Makarin (2020). Production Networks and War. *Available at SSRN 3969161*.

Lafrogne-Joussier, R., J. Martin, and I. Mejean (2022). Supply Shocks in Supply Chains: Evidence from the Early Lockdown in China. *IMF economic review*, 1–46.

Lagos, R. (2006). A Model of TFP. *The Review of Economic Studies 73*(4), 983–1007.

Liu, E. and A. Tsyvinski (2020). Dynamical Structure and Spectral Properties of Input-Output Networks. *NBER Working Paper 28178*.

Mira, A. N. (1964). Moscow: Miklukho-Maklai Ethnological Institute. *Department of Geodesy and Cartography, State Geological Committee of the Soviet Union*.

Miranda-Pinto, J., A. Silva, and E. R. Young (2022). Business cycle asymmetry and input-output structure: The role of firm-to-firm networks.

Panigrahi, P. P. (2021). Endogenous Spatial Production Networks: Theory, Estimation, and Evidence from Indian Firm-to-Firm Linkages. *CESifo Working Paper 9466*.

Peter, A. and C. Ruane (2022). The Aggregate Importance of Intermediate Input Substitutability. *Working Paper*.

Rauch, J. E. (1999). Networks versus Markets in International Trade. *Journal of international Economics 48*(1), 7–35.

Ravindran, S. and M. Shah (2020). Unintended Consequences of Lockdowns: Covid-19 and the Shadow Pandemic. *27562*.

Redding, S. J. and D. E. Weinstein (2020). Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES preferences. *The Quarterly Journal of Economics 135*(1).

Taschereau-Dumouchel, M. (2020). Cascades and Fluctuations in an Economy with an Endogenous Production Network. *Available at SSRN 3115854*.

Appendix for online publication only

A Appendix Tables

TABLE A1: Summary statistics

	Jan-March	April-June	July-September
		Panel A: 2019	
Number of sellers	135,849	131,996	133,897
Number of buyers	193,660	188,708	189,219
Total sales (mln. rupees)	962,688	908,361	1,036,831
Number of transactions	7,772,883	7,808,325	7,934,706
		Panel B: 2020	
	Jan-March	April-June	July-September
Number of sellers	113,121	69,171	86,696
Number of buyers	164,153	114,353	135,056
Total sales (mln. rupees)	811,755	369,645	775,478
Number of transactions	7,362,508	3,201,081	4,782,336

Notes: This table consists of two panels. Panel A contains information about the number of sellers, buyers, transactions, and total sales for the periods January-March, April-June, July-September for the year 2019. Panel B is the same as Panel A, but for 2020.

TABLE A2: Distribution of economic activity by industry and type of transaction

HS section	Sales share	Purchase share
Animals	1.5034	0.7723
Vegetables	15.2982	11.2945
Fats	2.2934	2.6251
Processed foods	4.2172	5.5548
Minerals	13.1241	10.2353
Chemicals	9.8288	9.0791
Plastics	13.1516	9.1410
Leather	0.1618	0.1677
Wood	2.5110	1.2130
Wood derivatives	1.0783	1.3598
Textiles	3.6342	6.4576
Clothing	1.3428	0.9107
Handicrafts	1.0190	1.9337
Jewelry	1.7005	1.4980
Metal	10.4473	12.1969
Machinery	10.9909	13.5771
Transport equipment	4.7124	8.4147
Surgical instrum.	1.4478	1.6478
Arms and ammo	0.0057	0.0095
Miscellaneous	1.2263	1.4936
Art	0.3043	0.4166
Type of transaction		
Within-state	72.6822	52.2224
Inter-state	23.2183	44.5151
Foreign	4.0994	3.2623

Notes: The table consists of an upper panel and a lower panel. In the upper panel, we show the share of sales to and purchases from our Indian state of analysis by industry (HS Section). In the lower panel, we show the share of sales to and purchases from our Indian state, by whether the buyer or seller is within the state, in another state of India, or abroad. Statistics were calculated using data for 2019.

TABLE A3: 2SLS, firm-level elasticity of substitution across (at least two) suppliers

	(1)	(2)	(3)	(4)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.2383	0.3381	0.4121	0.3688
	(0.1206)	(0.0627)	(0.1236)	(0.1146)
ϵ	0.7616	0.6618	0.5878	0.6311
Obs	851120	599918	599918	544819
K-PF	58.989	97.958	233.084	527.534
Seller IV	Y		Y	Y
Bilateral IV		Y	Y	Y
Buyer-month FE	Y	Y	Y	Y
Product-month FE	Y	Y	Y	Y
Buyer-product FE				Y
Seller-product FE				Y

Notes: 2SLS estimates are obtained from estimating Equation (11). The set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$. That is, a supplier s of buyer b is considered *common* if they traded during both the current and previous month. We only consider the cases when a buyer traded with at least two common suppliers in a given period. The first stage uses either bilateral or seller-level instruments, as pointed out by rows six and seven. Bilateral instruments correspond to Equation (14), while seller-level instruments correspond to Equation (13). The first row reports the estimates associated with changes in log relative unit values. Standard errors are two-way clustered at the origin and destination state level, and are reported in parentheses below each estimate. The third row reports the implied value for ϵ, which is 1 minus the estimate on the first row. The fourth row reports the number of observations. The fifth row reports the Kleibergen-Paap F statistic from the first stage. A product category is a 4-digit HS code and the treatment period is March-May 2020. The table contains four columns. Each column corresponds to different combinations of instruments and fixed effects, as pointed out by the last six rows. All specifications include the controls mentioned in the paper.

TABLE A4: Alternative specifications: 2SLS, firm-level elasticity of substitution across suppliers

	(1)	(2)	(3)	(4)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.5687	0.5476	0.9371	0.8063
	(0.2086)	(0.1818)	(0.3856)	(0.3305)
ϵ	0.4312	0.4523	0.0628	0.1936
Obs	879997	851483	1026381	993583
K-PF	37.629	121.309	42.335	87.990
HS digits	6	6	8	8
Seller IV	Y	Y	Y	Y
Bilateral IV	Y	Y	Y	Y
Buyer-month FE	Y	Y	Y	Y
Product-month FE	Y	Y	Y	Y
Buyer-product FE		Y		Y
Seller-product FE		Y		Y

Notes: 2SLS estimates are obtained from estimating Equation (11). The set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$. That is, a supplier s of buyer b is considered *common* if they also traded during both the current and previous month. In all specifications, the first stage uses both bilateral and seller-level instruments as pointed out in rows seven and eight. Bilateral instruments correspond to Equation (14), while seller-level instruments correspond to Equation (13). The first row reports the estimates associated with changes in relative unit values in logs. Standard errors are two-way clustered at the origin and destination state level, and are reported in parentheses below each estimate. The third row reports the implied value for ϵ, which is 1 minus the estimate on the first row. The fourth row reports the number of observations. The fifth row reports the Kleibergen-Paap F statistic from the first stage. A product category is either a 6-digit or 8-digit HS code, as pointed out by the sixth row, and the treatment period is March-May 2020. The table contains four columns. Each column corresponds to different combinations of HS codes and fixed effects, as pointed out by the last six rows. All specifications include the controls mentioned in the paper.

TABLE A5: Firm-level elasticities of substitution across suppliers, by HS section

Section	Name	OLS elast.	2SLS elast.
1	Animals	0.6892	0.1648
2	Vegetables	0.7799	0.7149
3	Fats	.	.
4	Processed foods	0.7125	0.1917
5	Minerals	0.8326	0.3974
6	Chemicals	0.7735	0.5828
7	Plastics	0.7179	0.9796
8	Leather	.	.
9	Wood	0.8728	0.6154
10	Wood derivatives	0.7812	0.8915
11	Textiles	0.8249	0.8103
12	Clothing	0.8232	0.3360
13	Handcrafts	0.6737	.
14	Jewelry	0.8104	1.3721
15	Metal	0.8145	0.8142
16	Machinery	0.6072	0.8691
17	Transport equipment	.	.
18	Surgical instruments	0.5954	0.3799
19	Arms and ammo	0.4140	.
20	Miscellaneous	0.6903	0.8383
21	Art	0.5514	0.1486

Notes: Each row corresponds to an industry, which is defined as an HS section. The second column contains the name of the industry. The third and fourth columns report the estimated elasticities by OLS and 2SLS from Equation (11). Both OLS and 2SLS estimators include product-by-month, buyer-by-month, buyer-by-product, and seller-by-product fixed effects. Standard errors are two-way clustered at both origin and destination states. All specifications include the controls mentioned in the paper. Elasticities were not reported if there was low statistical power or a weak first stage.

TABLE A6: 2SLS, firm-level elasticity of substitution across suppliers, by institutional quality

	Sellers' Institutional quality			Buyers' Institutional quality		
	Low	Mid	High	Low	Mid	High
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	-0.5912	0.2670	0.7686	-0.0435	0.9620	0.2796
	(0.2052)	(0.3111)	(0.3395)	(0.1298)	(0.4155)	(0.2615)
ϵ	1.5912	0.7329	0.2313	1.0435	0.0379	0.7203
Obs	642177	1351320	440377	1087166	1258625	479662
K-PF	42.526	336.600	37.385	45.178	10.831	194.983

Notes: The first row reports 2SLS estimates from Equation (11). For the 2SLS estimates, the set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$. That is, a supplier s of buyer b is considered *common* if they also traded during the previous period. The first stage uses both bilateral and seller-level instruments. Bilateral instruments correspond to Equation (14), while seller-level instruments correspond to Equation (13). Standard errors are two-way clustered at the origin and destination state level, and are reported in parentheses below each estimate. The third row reports the implied value for ϵ. The fifth row reports Kleibergen-Paap F statistics from the first stage. The first three columns report estimates conditional on the institutional quality of sellers' location, and the last three columns report estimates conditional on the institutional quality of buyers' location. The locations of sellers or buyers are categorized in terciles according to their institutional quality. Districts in the first tercile of institutional quality are categorized as *Low*, in the second tercile are categorized as *Medium*, and districts in the third tercile of institutional quality are categorized as *High*. This variable is measured as the average number of days a district's courts take to reach a verdict. All specifications include buyer-by-month and product-by-month fixed effects.

TABLE A7: Firm-level elasticity of substitution across suppliers, monthly vs. quarterly

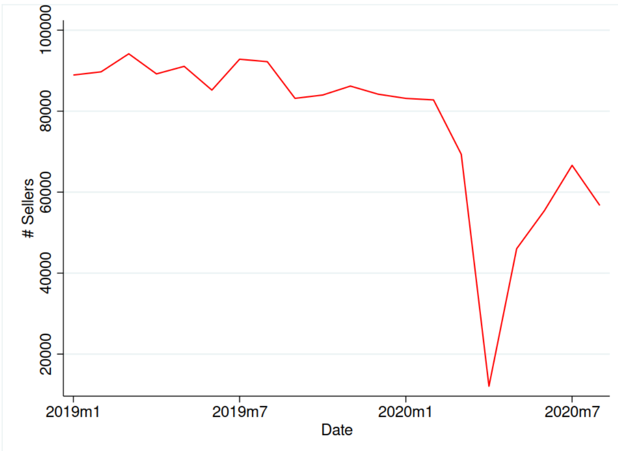
	Monthly		Quarterly	
	OLS	2SLS	OLS	2SLS
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.2171	0.4538	0.1436	0.2116
	(0.0133)	(0.5461)	(0.8564)	(0.7883)
ϵ	0.7828	0.5461	0.8564	0.7883
Obs	2028039	2028039	1518102	1518102
K-PF	.	143.413	.	11.501

Notes: The first row report the OLS estimates from Equation (11), and 2SLS estimates from Equation (11). For the 2SLS estimates, the set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$. That is, a supplier s of buyer b is considered *common* if they also traded during the previous period. The first stage uses both bilateral and seller-level instruments. Bilateral instruments correspond to Equation (14), while seller-level instruments correspond to Equation (13). Periods are either months or quarters depending on the frequency of the data. The first two columns are monthly estimates, and the last two columns are quarterly estimates. Standard errors are two-way clustered at the origin and destination state level, and are reported in parentheses below each estimate. The third row reports the implied value for ϵ. The fourth row reports the number of observations. The fifth row reports Kleibergen-Paap F statistics from the first stage. All specifications include buyer-by-period and product-by-period fixed effects.

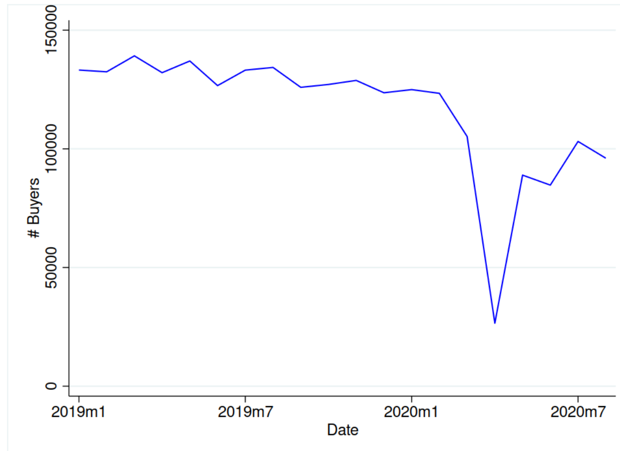
B Appendix Figures

Figure A1: Variation over time in aggregate outcomes

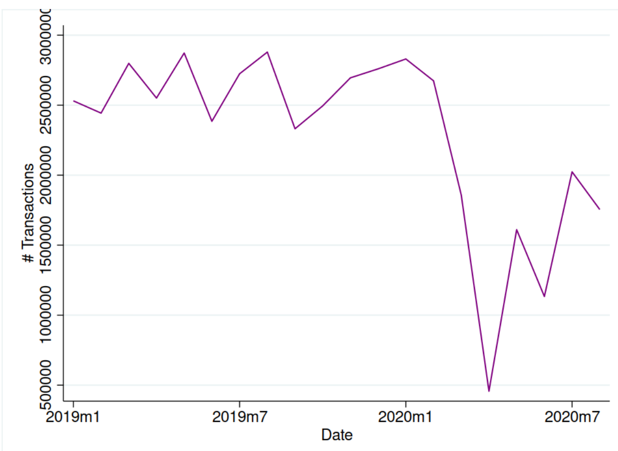
(a) Number of sellers



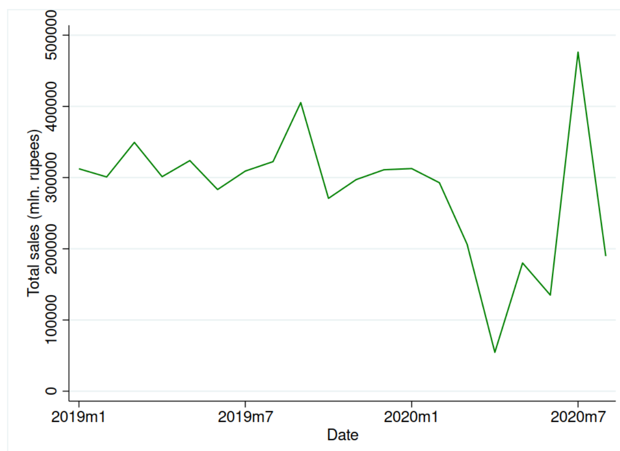
(b) Number of buyers



(c) Number of transactions



(d) Total sales

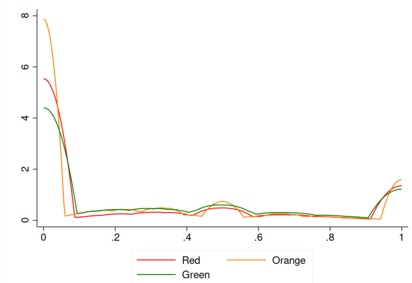


Notes: This figure consists of 4 panels. In each panel, the horizontal axis is time, and the vertical axis is a different aggregate outcome. In the first panel, we show the number of sellers that reported a transaction by period. In the the second panel, we show the number of buyers that reported a transaction by period. In the third panel, we show the number of transactions that were reported in a given period. In the fourth panel, we show the total sales for a given period.

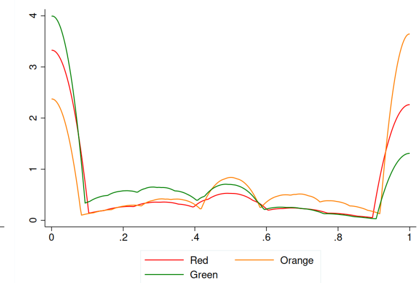
FIGURE A2: Distribution of links and sales across lockdown zones

Share distributions of colors

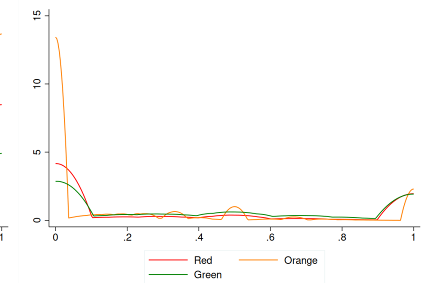
(a) Sellers in *Red*



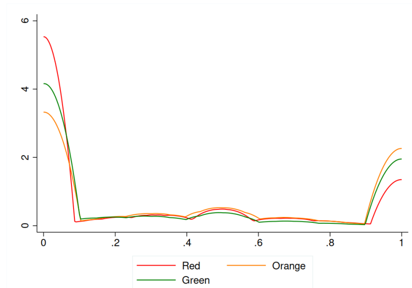
(b) Sellers in *Orange*



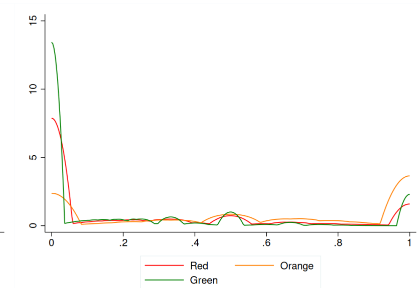
(c) Sellers in *Green*



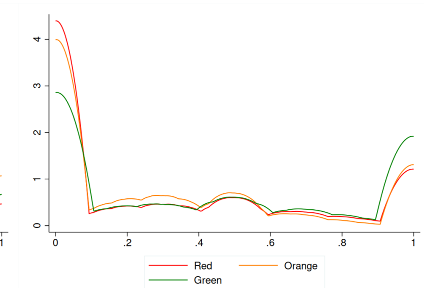
(d) Buyers in *Red*



(e) Buyers in *Orange*

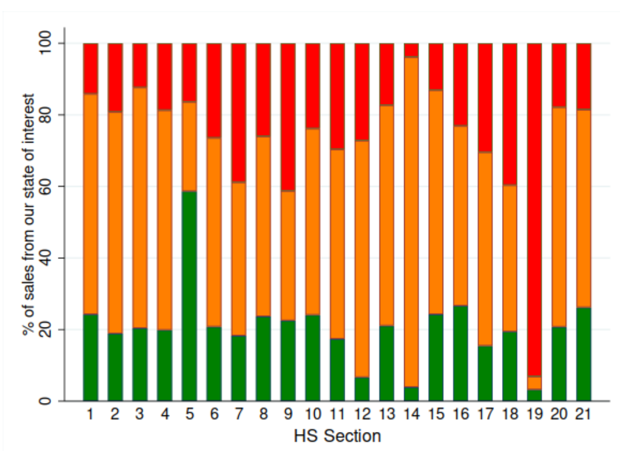


(f) Buyers in *Green*

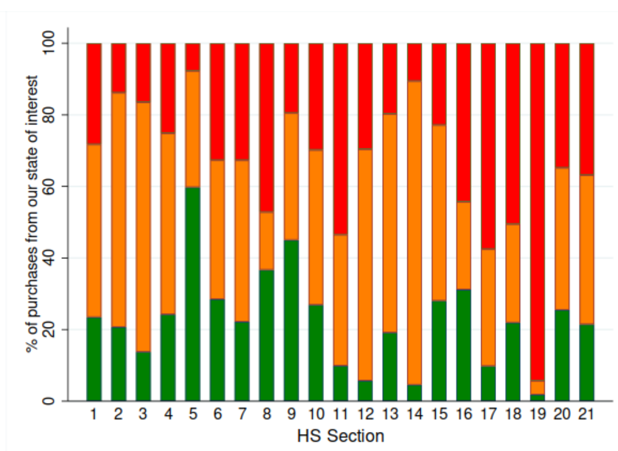


Share of sales and purchases by color zone of destination districts

(g) Sales



(h) Purchases



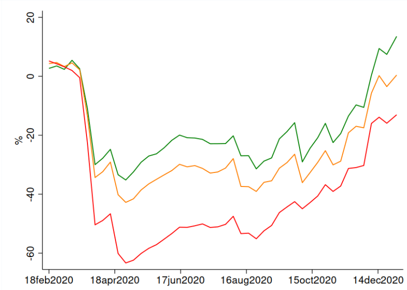
Notes: This figure comprises two sets of panels. The first six figures are the first panel, and the last two figures are the second panel. For the first panel, in the three upper sub-figures, each panel plots the distribution of the share of buyers located in *Red*, *Orange*, or *Green* districts. Each sub-figure corresponds to sellers located in their corresponding color district. In the middle three sub-figures, each sub-figure plots the distribution of the share of sellers located in *Red*, *Orange*, or *Green* districts. Each sub-figure corresponds to buyers located in their corresponding color district. The time period is April 2018 - February 2020. For the lower panel, on the left sub-figure, for each HS section (horizontal axis), we plot the share of total sales of firms located in our large Indian state by the zone of selling districts. In the lower right sub-figure, for each HS section (horizontal axis), we plot the share of total purchases of firms located in our large Indian state by the zone of buying districts. The time period for this data is 2019.

FIGURE A3: Google mobility trends by lockdown zone

(a) Retail and recreation

(b) Grocery and pharmacy

(c) Parks



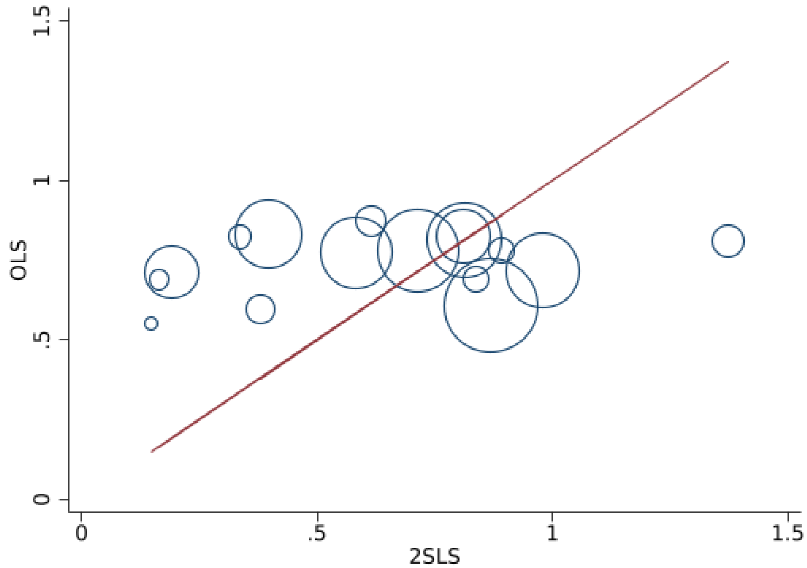
(d) Transit stations

(e) Workplaces

(f) Residential

Notes: These plots are constructed using *Google Mobility Trends* data, which shows how visits and length of stay at different places change compared to a baseline. The baseline is the median value, for the corresponding day of the week, during January 3rd - February 6th 2020. The raw data is at the daily frequency for each district in India. We collapse the data at a weekly frequency, and at the zone level. Each panel corresponds to mobility in different places.

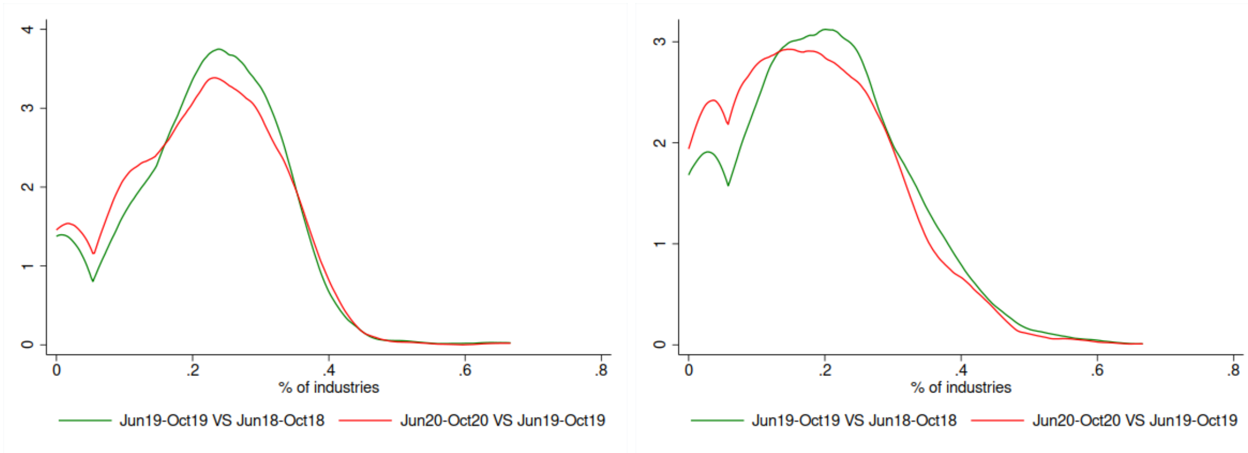
FIGURE A4: Elasticities ϵ by seller's industry



Notes: The vertical axis is the OLS estimate of ϵ, and the horizontal axis is the 2SLS estimate of ϵ. These estimates come from estimating Equation (11). For the 2SLS estimate, the set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$. That is, a supplier s of buyer b is considered *common* if they traded in both the current and previous month. The first stage uses both bilateral and seller-level instruments. Bilateral instruments correspond to Equation (14), and seller-level instruments correspond to Equation (13). An industry is an HS section. The size of each bubble is determined by total sales in the corresponding industry. See Table A5 for industry-specific numbers.

FIGURE A5: Change in Product Category Links, before and after lockdown

(a) Sellers (b) Buyers



Notes: The figure has two density plots. On the left, we study sellers; on the right, buyers. In the left plot, we show the distribution of the share of sellers that sold goods to a given product category in both periods t and $t-1$, where these periods are one year apart. In the right plot, we show the distribution of the share of buyers that purchased goods from a given product category in both periods t and $t-1$, where these periods are one year apart. Product categories are 4-digit HS codes. The green densities are for periods before Covid-19 lockdowns, where t is between June 2019 and October 2019, and $t-1$ is between June 2018 and October 2018. The red densities are for periods after Covid-10 lockdowns, where t is between June 2020 and October 2020, and $t-1$ is between June 2019 and October 2019.

C Data

Exposure variables. We construct two exposure variables at the firm level: $ED_{si,t}$ and $IM_{si,t}$. $ED_{si,t}$ is the exposure of firm s selling product i to global demand shocks in month t. $IM_{si,t}$ is the exposure of firm s selling product i to global supply shocks in month t. First, we construct these exposures by country, such that

$$ED_{si,x,t} = \left(\frac{Y_{si,x,0}}{\sum_{x'} Y_{si,x',0}}\right) X_{i,x,t}$$
$$IM_{si,m,t} = \left(\frac{Y_{si,m,0}}{\sum_{m'} Y_{si,m',0}}\right) M_{i,m,t},$$

where $Y_{si,x,0}$ is the value of goods of seller s of product i shipped to country x in 2018, $Y_{si,m,0}$ is the value of goods of seller s of product i shipped from country m in 2018, $X_{i,x,t}$ is the value of export demand from country x for product i in month t, excluding demand for Indian products, and $M_{i,m,t}$ is the value of import demand to country x for product i in month t, excluding demand for Indian products. We then construct our exposure variables as a weighted sum of these measures across countries, such that

$$ED_{si,t} = \sum_{x} \left(\frac{Y_{s,x,0}}{\sum_{x'} Y_{s,x',0}}\right) ED_{si,x,t}$$
$$IM_{si,t} = \sum_{m} \left(\frac{Y_{s,m,0}}{\sum_{m'} Y_{s,m',0}}\right) IM_{si,m,t},$$

where $Y_{s,x,0}$ is the value of goods of seller s shipped to country x in 2018, and $Y_{s,m,0}$ is the value of goods of seller s shipped from country m in 2018.

Labor and sales. Our firm-to-firm dataset lacks data on the number of employees and final sales. We obtain data on the number of employees and total sales from an external dataset for a subset of our firms. We then estimate an OLS regression of both labor and final sales on observable variables in our firm-to-firm dataset, store the OLS estimates, and use them to predict labor and final sales for all firms.

We scrape data on the number of employees and total sales from the website *IndiaMART*,[26] India's largest B2B digital platform. We scraped around 300,000-400,000 firm profiles, and then sent them to the tax authority to be matched with our firm-to-firm trade dataset. The matching

[26] https://www.indiamart.com/

procedure (conducted by the government) yielded 50,720 unique firms.

Each firm reports its number of employees and annual turnover (sales), both reported in brackets. The reported brackets for sales are: up to 50 Lakh, 50 Lakh-1 Crore, 1-2 Crore, 2-5 Crore, 5-10 Crore, 10-25 Crore, 25-50 Crore, 50-100 Crore, 100-500 Crore 500-1,000 Crore, 1,000-5,000 Crore, 5,000-10,000 Crore, more than 10,000 Crore. First, we convert each reported number into rupees, since sales in the trade dataset are reported in rupees.[27] Then, for each firm we assign the median value of its corresponding sales bracket. For the last bracket, we consider the upper bound to be 100,000 Crore. The reported brackets for labor are: up to 10 employees, 11-25, 26-50, 51-100, 101-500, 501-1000, 1001-2000, 2001-5000, and more than 5000 employees. For each firm, we assign the median value of its corresponding labor bracket. For the last bracket, we consider the upper bound to be 50,000 employees.

We then estimate the following OLS regressions:

$$\log(labor_n) = \alpha_0 + \alpha_1 \log(sales_n) + \alpha_2 \log(distance_n) + \epsilon_i^l$$
$$\log(final_n) = \beta_0 + \beta_1 \log(sales_n) + \beta_2 \log(distance_n) + \epsilon_i^f,$$

where $sales_n$ are total sales of intermediates of firm n, $distance_n$ is the average distance in kilometers of all firms' registered transactions, $labor_n$ is the number of employees constructed as previously explained, and $final_n$ is final sales. We constructed final sales by subtracting total intermediate sales from total sales, where we construct the former directly from our firm-to-firm dataset. In almost all cases, this difference was positive, which reassures that IndiaMART indeed reports total sales. Whenever the differences were negative, we input a value of 0, which implies that all firm's sales are of intermediates.

We obtain the following estimated elasticities: $(\widehat{\alpha}_0, \widehat{\alpha}_1, \widehat{\alpha}_2) = (-2.1138, 0.2502, 0.2853)$, and $(\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2) = (9.8848, 0.3665, 0.4227)$. They are significant at the 1% confidence level. We then use these estimates to predict labor and final sales to all firms in our dataset.

Dijkstra algorithm We now list the steps of the Dijkstra algorithm we used to construct our seller-buyer-level instruments. We obtained a set of *shapefiles* of district administrative boundaries for India according to India's 2011 census. We reprojected the shapefiles into an *Asian/South Equidistance Conic* projection, which is the projection that best preserves the distance measurements. Once shapefiles are reprojected, the objective is to construct a transportation network between Indian districts.

[27] 100,000 rupees = 1 Lakh; and 10,000,000 rupees = 1 Crore.

First, we obtain the centroid of each district in India. Then, we construct a network structure according to the set of centroids. There are many ways to construct a network, so we need to take a stance on how to form the connections between centroids. For each centroid, we generate connections to the k closest centroids according to Euclidean distances.[28] We follow Fajgelbaum and Schaal (2020) and consider $k = 8$ such that we consider the main cardinal directions (i.e. north, south, east, west, north-east, south-east, north-west, south-west).

We now run the Dijkstra algorithm. For all district pairs, the algorithm provides us with the list of all districts that comprise the route between the district pair, and the distance of each leg that comprise the route. Using the names of the districts, we use the lockdown data to assign a lockdown color to each district along the route, and obtain our seller-buyer-level instruments. Our first instrument is the share of districts in a route that are *Red*, *Orange*, or *Green*. When calculating these shares, we rule out the zone where the buyer resides so we do not consider demand-side shocks in our instrument. Using the distance of each leg, our second instrument is the share of meters of the route that are *Red*, *Orange*, or *Green*. We consider a leg to be of color $x = Red, Orange, Green$ whenever the origin district was of color x. In this case, we also ignore the color of the district where the buyer resides.

D DERIVATIONS

D.1 Estimation of firm-level elasticities of substitution across suppliers

In this section, we describe the steps to derive the firm-level elasticity of substitution across suppliers for the same product. First, we describe the model and the equations we take to the data. Second, explain how we construct price indices we need to estimate this elasticity. Third, we describe how we account for the entry and exit of suppliers for the estimation. Finally, we explain how we construct the seller-level and seller-buyer-level instruments we use to causally estimate our elasticity.

[28]Consider the set of nodes Φ, where $K \equiv |\Phi|$ is the number of nodes. The number of connections per node k could range from 0 up to K, where each represents extreme cases of network formation. $k = 0$ is a network without connections, so it is not possible to run a Dijkstra algorithm since it is not possible to go from one node to another. $k = K$ is a fully-connected network, where all nodes are connected with each other. Running a Dijkstra algorithm on this scenario is trivial since the shortest distance between any pair of nodes is their connection itself. Therefore, a feasible number of connections per node must be $k \in (0, K)$.

D.1.1. Expression to estimate firm-level elasticities of substitution across suppliers

A firm b selling product $j \in F$ maximizes profits subject to its technology and to a CES bundle of intermediate inputs:

$$\begin{aligned}
\max \quad & p_{bj}y_{bj} - w_{bj}l_{bj} - \sum_i \sum_s p_{si,bj}x_{si,bj} \\
& s.t. \\
& y_{bj} = A_b \left(w_{bl} \left(l_{bj} \right)^{\frac{\alpha-1}{\alpha}} + (1-w_{bl}) \left(x_{bj} \right)^{\frac{\alpha-1}{\alpha}} \right)^{\frac{\alpha}{\alpha-1}}, \\
& x_{bj} = \left(\sum_i w_{i,bj}^{\frac{1}{\zeta}} x_{i,bj}^{\frac{\zeta-1}{\zeta}} \right)^{\frac{\zeta}{\zeta-1}}, \\
& x_{i,bj} = \left(\sum_s \mu_{si,bj}^{\frac{1}{\epsilon}} x_{si,bj}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}
\end{aligned}$$

The first order condition with respect to $x_{si,bj}$ is

$$\begin{aligned}
\left[x_{si,bj}\right] : & p_{bj} \left(\frac{\alpha}{\alpha-1} \right) y_{bj} \left(\Theta^1_{bj} \right)^{-1} (1-w_{bl}) \left(\frac{\alpha-1}{\alpha} \right) x_{bj}^{\frac{\alpha-1}{\alpha}-1} \\
& \left(\frac{\zeta}{\zeta-1} \right) x_{bj} \left(\Theta^2_{bj} \right)^{-1} w_{i,j} \left(\frac{\zeta}{\zeta-1} \right) x_{i,bj}^{\frac{\zeta-1}{\zeta}-1} \\
& \left(\frac{\epsilon}{\epsilon-1} \right) x_{i,bj} \left(\Theta^3_{i,bj} \right)^{-1} \mu_{si,bj}^{\frac{1}{\epsilon}} \left(\frac{\epsilon-1}{\epsilon} \right) x_{si,bj}^{\frac{\epsilon-1}{\epsilon}-1} = p_{si,bj}, \\
= & p_{bj}y_{bj} \left(\Theta^1_{bj} \right)^{-1} (1-w_{bl}) x_{bj}^{\frac{\alpha-1}{\alpha}} \\
& \left(\Theta^2_{bj} \right)^{-1} w_{i,j} x_{i,bj}^{\frac{\zeta-1}{\zeta}} \\
& \left(\Theta^3_{i,bj} \right)^{-1} \mu_{si,bj}^{\frac{1}{\epsilon}} x_{si,bj}^{-\frac{1}{\epsilon}} = p_{si,bj},
\end{aligned}$$

where $\{\Theta^1_{bj}, \Theta^2_{bj}, \Theta^3_{i,bj}\}$ are composite terms that cancel out in the next steps. Now, consider the first order conditions with respect to $x_{si,bj}$ and $x_{s'i,bj}$ and divide them, such that

$$
\begin{aligned}
\frac{\mu_{si,bj}^{\frac{1}{\epsilon}} x_{si,bj}^{\frac{-1}{\epsilon}}}{\mu_{s'i,bj}^{\frac{1}{\epsilon}} x_{s'i,bj}^{-\frac{1}{\epsilon}}} &= \frac{p_{si,bj}}{p_{s'i,bj}}, \\
\frac{x_{si,bj}^{\frac{-1}{\epsilon}} p_{si,bj}^{\frac{-1}{\epsilon}}}{x_{s'i,bj}^{-\frac{1}{\epsilon}} p_{s'i,bj}^{\frac{-1}{\epsilon}}} &= \frac{p_{si,bj}^{1-\frac{1}{\epsilon}} \mu_{si,bj}^{-\frac{1}{\epsilon}}}{p_{s'i,bj}^{1-\frac{1}{\epsilon}} \mu_{s'i,bj}^{-\frac{1}{\epsilon}}}, \\
\left(x_{si,bj} p_{si,bj}\right)^{-\frac{1}{\epsilon}} \left(p_{s'i,bj}^{\frac{\epsilon-1}{\epsilon}} \mu_{s'i,bj}^{-\frac{1}{\epsilon}}\right) &= p_{si,bj}^{\frac{\epsilon-1}{\epsilon}} \mu_{si,bj}^{-\frac{1}{\epsilon}} \left(x_{s'i,bj} p_{si,bj}\right)^{-\frac{1}{\epsilon}}, \\
\left(x_{si,bj} p_{si,bj}\right) \left(p_{s'i,bj}^{1-\epsilon} \mu_{s'i,bj}\right) &= p_{si,bj}^{1-\epsilon} \mu_{si,bj} \left(x_{s'i,bj} p_{si,bj}\right), \\
\left(PM_{si,bj}\right) \left(p_{s'i,bj}^{1-\epsilon} \mu_{s'i,bj}\right) &= p_{si,bj}^{1-\epsilon} \mu_{si,bj} \left(PM_{s'i,bj}\right), \\
\left(PM_{si,bj}\right) \sum_{s'} \left(p_{s'i,bj}^{1-\epsilon} \mu_{s'i,bj}\right) &= p_{si,bj}^{1-\epsilon} \mu_{si,bj} \sum_{s'} \left(PM_{s'i,bj}\right), \\
\left(PM_{si,bj}\right) p_{i,bj}^{1-\epsilon} &= p_{si,bj}^{1-\epsilon} \mu_{si,bj} PM_{i,bj}, \\
\frac{PM_{si,bj}}{PM_{i,bj}} &= \left(\frac{p_{si,bj}}{p_{i,bj}} \mu_{si,bj}^{\frac{1}{1-\epsilon}}\right)^{1-\epsilon}, \\
\log\left(\frac{PM_{si,bj}}{PM_{i,bj}}\right) &= (1-\epsilon)\log\left(\frac{p_{si,bj}}{p_{i,bj}}\right) + \log\left(\mu_{si,bj}\right),
\end{aligned}
$$

where $PM_{si,bj} \equiv p_{si,bj} x_{si,bj}$, $p_{i,bj}^{1-\epsilon} \equiv \sum_{s'} p_{s'i,bj}^{1-\epsilon} \mu_{s'i,bj}$, and $PM_{i,bj} \equiv \sum_{s'} PM_{s'i,bj}$. This is the derivation of Equation (9).

D.1.2. Constructing price indices

In this section, we derive the expressions that allow us to construct price indexes based on observable data. First, go back to the derivation in Appendix D.1, where

$$\left(PM_{si,bj}\right) p_{i,bj}^{1-\epsilon} = p_{si,bj}^{1-\epsilon} \mu_{si,bj} PM_{i,bj}.$$

In the data, we observe prices and expenditures over time, so we introduce a time dimension such that

$$\left(PM_{si,bj,t}\right) p_{i,bj,t}^{1-\epsilon} = p_{si,bj,t}^{1-\epsilon} \mu_{si,bj,t} PM_{i,bj,t},$$

where t is a month. We can now express this equation in changes, such that

$$\left(\widehat{PM}_{si,bj,t}\right) \widehat{p}_{i,bj,t}^{1-\epsilon} = \widehat{p}_{si,bj,t}^{1-\epsilon} \widehat{\mu}_{si,bj,t} \widehat{PM}_{i,bj,t},$$

where $\widehat{x}_t \equiv \frac{x_t}{x_{t-1}}$. Our objective is for $\widehat{p}_{i,bj,t}$ not to depend on $\widehat{\mu}_{si,bj,t}$, which are not observable. To do this, we rely on Redding and Weinstein (2020). The key assumption is that the overall importance of a product category in a buyer's input use is time-invariant. Concretely, the geometric mean of $\mu_{si,bj,t}$ across common sellers is constant. From the maximization problem of the firm, we obtain the following expression for the CES price index at the buyer level:

$$p_{i,bj,t} = \left(\sum_{s \in \Omega_{i,bj,t}} \mu_{si,bj,t} p_{si,bj,t}^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}},$$

where $\Omega_{i,bj,t}$ is the set of all sellers that provided to buyer b in time t. We apply Shephard's Lemma to this CES price function, which in turn yields an expression for expenditure share:

$$s_{si,bj,t} = \frac{\mu_{si,bj,t} p_{si,bj,t}^{1-\epsilon}}{p_{i,bj,t}^{1-\epsilon}},$$

where $s_{si,bj,t} \equiv \frac{PM_{si,bj,t}}{\sum_{s \in \Omega_{i,bj,t}} PM_{si,bj,t}}$. We can then rewrite this expression such that

$$p_{i,bj,t} = p_{si,bj,t} \left(\frac{\mu_{si,bj,t}}{s_{si,bj,t}} \right)^{\frac{1}{1-\epsilon}}, \forall s \in \Omega_{i,bj,t}.$$

This expression in changes is

$$\widehat{p}_{i,bj,t} = \widehat{p}_{si,bj,t} \left(\frac{\widehat{\mu}_{si,bj,t}}{\widehat{s}_{si,bj,t}} \right)^{\frac{1}{1-\epsilon}}.$$

Now, common suppliers for a buyer b in time t is the set of suppliers $\Omega^*_{i,bj,t}$ that sold to buyer b in the current and previous period (i.e. $\Omega^*_{i,bj,t} \equiv \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$), where $N^*_{i,bj,t} \equiv \left|\Omega^*_{i,bj,t}\right|$ is the number of common sellers for buyer b in time t. We now apply a geometric mean to this expression, such that

$$
\begin{aligned}
\widehat{p}_{i,bj,t}^{N^*_{i,bj,t}} &= \prod_{s=1}^{N^*_{i,bj,t}} \left\{ \widehat{p}_{si,bj,t} \left(\frac{\widehat{\mu}_{si,bj,t}}{\widehat{s}_{si,bj,t}} \right)^{\frac{1}{1-\epsilon}} \right\}, \\
\widehat{p}_{i,bj,t}^{N^*_{i,bj,t}} &= \prod_{s=1}^{N^*_{i,bj,t}} \widehat{p}_{si,bj,t} \prod_{s=1}^{N^*_{i,bj,t}} \widehat{\mu}_{si,bj,t}^{\frac{1}{1-\epsilon}} \prod_{s=1}^{N^*_{i,bj,t}} \widehat{s}_{si,bj,t}^{\frac{1}{\epsilon-1}}, \\
\widehat{p}_{i,bj,t} &= \prod_{s=1}^{N^*_{i,bj,t}} \widehat{p}_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}} \left(\prod_{s=1}^{N^*_{i,bj,t}} \widehat{\mu}_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}} \right)^{\frac{1}{1-\epsilon}} \prod_{s=1}^{N^*_{i,bj,t}} \left(\widehat{s}_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}} \right)^{\frac{1}{\epsilon-1}}, \\
\widehat{p}_{i,bj,t} &= \widehat{\widetilde{p}}_{i,bj,t} \widehat{\widetilde{s}}_{i,bj,t}^{\frac{1}{\epsilon-1}} \left(\prod_{s=1}^{N^*_{i,bj,t}} \widehat{\mu}_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}} \right)^{\frac{1}{1-\epsilon}}.
\end{aligned}
$$

We now formally state the assumption we require to move forward, which is

$$
\widetilde{\mu}_{i,bj,t} = \prod_{s=1}^{N^*_{i,bj,t}} \mu_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}} = \prod_{s=1}^{N^*_{i,bj,t}} \mu_{si,bj,t-1}^{\frac{1}{N^*_{i,bj,t}}} = \widetilde{\mu}_{i,bj,t-1}.
$$

Then, the last term of our expression is

$$
\begin{aligned}
\prod_{s=1}^{N^*_{i,bj,t}} \widehat{\mu}_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}} &= \prod_{s=1}^{N^*_{i,bj,t}} \left(\frac{\mu_{si,bj,t}}{\mu_{si,bj,t-1}} \right)^{\frac{1}{N^*_{i,bj,t}}}, \\
&= \frac{\prod_{s=1}^{N^*_{i,bj,t}} \mu_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}}}{\prod_{s=1}^{N^*_{i,bj,t}} \mu_{si,bj,t-1}^{\frac{1}{N^*_{i,bj,t}}}}, \\
&= \frac{\widetilde{\mu}_{i,bj,t}}{\widetilde{\mu}_{i,bj,t-1}}, \\
&= 1.
\end{aligned}
$$

So our final expression boils down to

$$
\widehat{p}_{i,bj,t}^{1-\epsilon} = \frac{\widehat{\widetilde{p}}_{i,bj,t}^{1-\epsilon}}{\widehat{\widetilde{s}}_{i,bj,t}},
$$

where $\widetilde{p}_{i,bj,t} \equiv \prod_s p_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}}$ is a geometric mean of unit values across common suppliers, and $\widetilde{s}_{i,bj,t} \equiv$

$\prod_s s_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}}$ is a geometric mean of expenditure shares across common suppliers. Notice that we have reached to our objective, since now $\widehat{\widetilde{p}}_{i,bj,t}$ is independent of $\mu_{si,bj,t}$. Finally, the expression we take to the data is

$$
\begin{aligned}
\left(\widehat{PM}_{si,bj,t}\right)\widehat{p}_{i,bj,t}^{1-\epsilon} &= \widehat{p}_{si,bj,t}^{1-\epsilon}\widehat{\mu}_{si,bj,t}\widehat{PM}_{i,bj,t}, \\
\left(\widehat{PM}_{si,bj,t}\right)\widehat{\widetilde{p}}_{i,bj,t}^{1-\epsilon}\widehat{\widetilde{s}}_{i,bj,t}^{-1} &= \widehat{p}_{si,bj,t}^{1-\epsilon}\widehat{\mu}_{si,bj,t}\widehat{PM}_{i,bj,t}, \\
\frac{\widehat{PM}_{si,bj,t}}{\widehat{PM}_{i,bj,t}} &= \left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right)^{1-\epsilon}\left(\widehat{\widetilde{s}}_{i,bj,t}\widehat{\mu}_{si,bj,t}\right), \\
\log\left(\frac{\widehat{PM}_{si,bj,t}}{\widehat{PM}_{i,bj,t}}\right) &= (1-\epsilon)\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right)+\log\left(\widehat{\widetilde{s}}_{i,bj,t}\widehat{\mu}_{si,bj,t}\right), \\
\log\left(\frac{\widehat{PM}_{si,bj,t}}{\widehat{PM}_{i,bj,t}}\right) &= (1-\epsilon)\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right)+\log\left(\widehat{\widetilde{s}}_{i,bj,t}\right)+\log\left(\widehat{\mu}_{si,bj,t}\right).
\end{aligned}
$$

D.1.3. Addressing the entry and exit of suppliers

In this section, we explain how we address the fact that seller and buyer matches do not happen in every period (i.e. entry and exit of sellers). The concern is that not taking into account the fact that sellers and buyers do not trade in every period could induce a bias in the estimation of ϵ. We address this by including a correction term by Feenstra (1994) in our regressions. First, notice we can write down the expenditure share as

$$s_{si,bj,t} \equiv \lambda_{i,bj,t}s^*_{si,bj,t},$$

where $\lambda_{i,bj,t}$ is the *Feenstra* correction term, and $s^*_{si,bj,t}$ is the expenditure share with respect to total expenditure on common suppliers. Notice that these terms are constructed as

$$
\begin{aligned}
s_{si,bj,t} &\equiv \frac{PM_{si,bj,t}}{\sum_{s\in\Omega_{i,bj,t}} PM_{si,bj,t}}, \\
\lambda_{i,bj,t} &\equiv \frac{\sum_{s\in\Omega^*_{i,bj,t}} PM_{si,bj,t}}{\sum_{s\in\Omega_{i,bj,t}} PM_{si,bj,t}}, \\
s^*_{si,bj,t} &\equiv \frac{PM_{si,bj,t}}{\sum_{s\in\Omega^*_{i,bj,t}} PM_{si,bj,t}}.
\end{aligned}
$$

In changes, the expression for expenditure shares is

$$\widehat{s}_{si,bj,t} = \widehat{\lambda}_{i,bj,t}\widehat{s}^*_{si,bj,t}.$$

Then, the geometric mean for expenditure shares is

$$\begin{aligned}\widehat{\widetilde{s}}_{i,bj,t} &= \prod_{s=1}^{N^*_{i,bj,t}} \widehat{s}_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}}, \\ &= \prod_{s=1}^{N^*_{i,bj,t}} \left(\widehat{\lambda}_{i,bj,t}\widehat{s}^*_{si,bj,t}\right)^{\frac{1}{N^*_{i,bj,t}}}, \\ &= \widehat{\lambda}_{i,bj,t} \prod_{s=1}^{N^*_{i,bj,t}} \left(\widehat{s}^*_{si,bj,t}\right)^{\frac{1}{N^*_{i,bj,t}}}, \\ &\widehat{\lambda}_{i,bj,t}\widehat{\widetilde{s}}^*_{i,bj,t}.\end{aligned}$$

So the final expression we take to the data is

$$\begin{aligned}\log\left(\frac{\widehat{PM}_{si,bj,t}}{\widehat{\widetilde{PM}}_{i,bj,t}}\right) &= (1-\epsilon)\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right) + \log\left(\widehat{\widetilde{s}}_{i,bj,t}\right) + \log\left(\widehat{\mu}_{si,bj,t}\right), \\ &= (1-\epsilon)\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right) + \log\left(\widehat{\lambda}_{i,bj,t}\widehat{\widetilde{s}}^*_{i,bj,t}\right) + \log\left(\widehat{\mu}_{si,bj,t}\right), \\ &= (1-\epsilon)\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right) + \log\left(\widehat{\lambda}_{i,bj,t}\right) + \log\left(\widehat{\widetilde{s}}^*_{i,bj,t}\right) + \log\left(\widehat{\mu}_{si,bj,t}\right).\end{aligned}$$

D.1.4. Addressing endogeneity concerns

The equation from the previous section is what we take to the data. Nevertheless, there are further endogeneity issues that would contaminate our estimates for ϵ. In particular, Covid lockdowns could have also induced changes in demand, which in turn would bias our estimates. For example, if Covid shocks also induce negative demand shocks, our estimates would then be biased upwards. In this section, we derive our instruments. First, we consider non-arbitrage in shipping, so prices at the origin and destination between sellers and suppliers are related as

$$p_{si,bj,t} = p_{si,t}\tau_{sb,t},$$

where $p_{si,t}$ is the marginal cost (MC) of production of good i for seller s in month t, $\tau_{sb,t}$ is the iceberg cost of transporting the good from seller s to buyer b in month t. Now, we can then express this in changes, such that

$$\widehat{p}_{si,bj,t} = \widehat{p}_{si,t}\widehat{\tau}_{sb,t}.$$

In logarithms, we have

$$\log\left(\widehat{p}_{si,bj,t}\right) = \log\left(\widehat{p}_{si,t}\right) + \log\left(\widehat{\tau}_{sb,t}\right).$$

These two components of price imply two instruments. First, our seller-level instrument uses variation in MC at the seller-product level due to lockdown measures at the seller's district. To isolate variation in marginal costs driven by seller's lockdown zone, we interact the lockdown indicator ($Lock_t$) which takes the value 1 between March and May with indicator variables Red_{o_s} and $Orange_{o_s}$ that equal 1 whenever seller s was located in a district o that was either *Red* or *Orange* during the lockdown. Then, our excluded instruments are

$$log(\widehat{p}_{si,t}) = \beta^{R,p} Red_{o_{(s)}} Lock_t + \beta^{O,p} Orange_{o_{(s)}} Lock_t + \nu^{p}_{si,t}.$$

Now we explain how we construct the instrument at the seller-buyer level. We have to take a stance about the functional form of the trade cost $\tau_{sb,t}$. We assume that trade costs are proportional to the travel time of the transportation of intermediate inputs, such that

$$\tau_{sb,t} = TravelTime^{\sigma}_{sb,t}.$$

If we express this in changes, we get

$$\widehat{\tau}_{sb,t} = \widehat{TravelTime}^{\sigma}_{sb,t}.$$

We exploit variation from the Covid-19 lockdown, which induced exogenous variation in the travel time between location pairs of sellers and buyers. Given this, we assume the following difference-in-differences setup for travel time:

$$\log\left(\widehat{TravelTime}_{sb,t}\right) = \beta^{R,T} Red_{o_{(s)}d_{(b)}} Lock_t + \beta^{O,T} Orange_{o_{(s)}d_{(b)}} Lock_t + \nu^{T}_{sb,t},$$

where $Red_{o_{(s)}d_{(b)}}$ and $Orange_{o_{(s)}d_{(b)}}$ are the share of the number of districts or of distance designated as *Red* and *Orange*, respectively, along the route between seller s and buyer b. We constructed these variables using Dijkstra algorithms. Further details about this are in Appendix C. Combining the expression for changes in travel time due to the lockdown and trade costs, we get the following expression for our seller-buyer-level excluded instruments

$$log(\widehat{\tau}_{sb,t}) = \beta^{R,\tau} Red_{o_{(s)}d_{(b)}} Lock_t + \beta^{O,\tau} Orange_{o_{(s)}d_{(b)}} Lock_t + \nu^{\tau}_{sb,t},$$

where $\beta^{R,\tau} \equiv \sigma\beta^{R,T}$, $\beta^{O,\tau} \equiv \sigma\beta^{O,T}$, and $\nu^{\tau}_{si,bj,t} \equiv \sigma\nu^{T}_{si,bj,t}$. Together, all excluded instruments are such that

$$\begin{aligned}\log\left(\widehat{p}_{si,bj,t}\right) = {} & \beta^{R,p} Red_{o_{(s)}} Lock_t + \beta^{O,p} Orange_{o_{(s)}} Lock_t + \\ & \beta^{R,\tau} Red_{o_{(s)}d_{(b)}} Lock_t + \beta^{O,\tau} Orange_{o_{(s)}d_{(b)}} Lock_t + \nu_{si,bj,t},\end{aligned}$$

where $\nu_{si,bj,t} \equiv \nu^{p}_{si,t} + \nu^{\tau}_{sb,t}$.

D.2 Estimation of firm-level elasticities of substitution across products

In this section, we describe the steps to derive the firm-level elasticity of substitution across products. First, we describe the model and the equations we take to the data. Second, we describe how we construct price indices we need to estimate this elasticity. Finally, we describe the instrument we use to causally estimate our elasticity.

D.2.1. Expressions to estimate firm-level elasticities of substitution across products

We rewrite the initial maximization problem, so

$$\begin{aligned}\max \quad & p_{bj}y_{bj} - w_{bj}l_{bj} - \sum_i p_{i,bj}x_{i,bj} \\ s.t. \quad & \\ y_{bj} &= A_b\left(w_{bl}\left(l_{bj}\right)^{\frac{\alpha-1}{\alpha}} + (1-w_{bl})\left(x_{bj}\right)^{\frac{\alpha-1}{\alpha}}\right)^{\frac{\alpha}{\alpha-1}}, \\ x_{bj} &= \left(\sum_i^I w_{i,bj}^{\frac{1}{\zeta}} x_{i,bj}^{\frac{\zeta-1}{\zeta}}\right)^{\frac{\zeta}{\zeta-1}}, \\ p_{i,bj} &= \left(\sum_s \mu_{si,bj} p_{si,bj}^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}\end{aligned}$$

The first order condition with respect to $x_{i,bj}$ is

$$\begin{aligned}
\left[x_{i,bj}\right] : & p_{bj}\left(\frac{\alpha}{\alpha-1}\right) y_{bj}\left(\Theta^4_{bj}\right)^{-1}(1-w_{bl})\left(\frac{\alpha-1}{\alpha}\right) x_{bj}^{\frac{\alpha-1}{\alpha}-1} \\
& \left(\frac{\zeta}{\zeta-1}\right) x_{bj}\left(\Theta^5_{bj}\right)^{-1} w_{i,bj}^{\frac{1}{\zeta}}\left(\frac{\zeta}{\zeta-1}\right) x_{i,bj}^{\frac{\zeta-1}{\zeta}-1} = p_{i,bj}, p_{i,bj} \\
& = p_{bj} y_{bj}\left(\Theta^4_{bj}\right)^{-1}(1-w_{bl}) x_{bj}^{\frac{\alpha-1}{\alpha}} \\
& \left(\Theta^5_{bj}\right)^{-1} w_{i,bj}^{\frac{1}{\zeta}} x_{i,bj}^{\frac{-1}{\zeta}},
\end{aligned}$$

where $\{\Theta^4_{bj}, \Theta^5_{bj}\}$ are composite terms that cancel out in the next steps. Now, consider the same first-order conditions with respect to $x_{i',bj}$ and divide them, such that

$$\begin{aligned}
\frac{w_{i,bj}^{\frac{1}{\zeta}} x_{i,bj}^{\frac{-1}{\zeta}}}{w_{i',bj}^{\frac{1}{\zeta}} x_{i',bj}^{\frac{-1}{\zeta}}} &= \frac{p_{i,bj}}{p_{i',bj}}, \\
\frac{w_{i,bj}^{\frac{1}{\zeta}} x_{i,bj}^{\frac{-1}{\zeta}} p_{i,bj}^{-\frac{1}{\zeta}}}{w_{i',bj}^{\frac{1}{\zeta}} x_{i',bj}^{\frac{-1}{\zeta}} p_{i',bj}^{-\frac{1}{\zeta}}} &= \frac{p_{i,bj} p_{i,bj}^{-\frac{1}{\zeta}}}{p_{i',bj} p_{i',bj}^{-\frac{1}{\zeta}}}, \\
\frac{w_{i',bj}\left(x_{i,bj} p_{i,bj}\right)}{w_{i,bj}\left(x_{i',bj} p_{i',bj}\right)} &= \frac{p_{i,bj}^{1-\zeta}}{p_{i',bj}^{1-\zeta}}, \\
PM_{i,bj}\left(w_{i',bj} p_{i',bj}^{1-\zeta}\right) &= PM_{i',bj}\left(w_{i,bj} p_{i,bj}^{1-\zeta}\right), \\
\sum_{i'} PM_{i,bj}\left(w_{i',bj} p_{i',bj}^{1-\zeta}\right) &= \sum_{i'} PM_{i',bj}\left(w_{i,bj} p_{i,bj}^{1-\zeta}\right), \\
PM_{i,bj} \sum_{i'} w_{i',bj} p_{i',bj}^{1-\zeta} &= w_{i,bj} p_{i,bj}^{1-\zeta} \sum_{i'} PM_{i',bj}, \\
PM_{i,bj} p_{bj}^{1-\zeta} &= w_{i,bj} p_{i,bj}^{1-\zeta} PM_{bj}, \\
\frac{PM_{i,bj}}{PM_{bj}} &= \frac{w_{i,bj} p_{i,bj}^{1-\zeta}}{p_{bj}^{1-\zeta}}, \\
\frac{PM_{i,bj}}{PM_{bj}} &= \left(w_{i,bj}^{\frac{1}{1-\zeta}} \frac{p_{i,bj}}{p_{bj}}\right)^{1-\zeta}, \\
\log\left(\frac{PM_{i,bj}}{PM_{bj}}\right) &= (1-\zeta)\log\left(\frac{p_{i,bj}}{p_{bj}}\right) + \log\left(w_{i,bj}\right),
\end{aligned}$$

where $PM_{bj} \equiv \sum_i PM_{i,bj}$, and $p_{bj} = \left(\sum_i w_{i,bj} p_{i,bj}^{1-\zeta}\right)^{\frac{1}{1-\zeta}}$. As we did for the estimation of the elasticity of substitution across suppliers, we introduce a time dimension, apply Shephard's lemma to this CES price function, and also assume that the overall importance of the composite intermediates is time-invariant, so

$$
\begin{aligned}
s_{i,bj,t} &= \frac{w_{i,bj,t} p_{i,bj,t}^{1-\zeta}}{p_{bj,t}^{1-\zeta}}, \\
p_{bj,t} &= p_{i,bj,t} \left(\frac{w_{i,bj,t}}{s_{i,bj,t}} \right)^{\frac{1}{1-\zeta}}, \\
\widehat{p}_{bj,t} &= \widehat{p}_{i,bj,t} \left(\frac{\widehat{w}_{i,bj,t}}{\widehat{s}_{i,bj,t}} \right)^{\frac{1}{1-\zeta}}, \\
\widehat{p}_{bj,t}^{N_{bj,t}} &= \prod_{i=1}^{N_{bj,t}} \widehat{p}_{i,bj,t} \left(\frac{\widehat{w}_{i,bj,t}}{\widehat{s}_{i,bj,t}} \right)^{\frac{1}{1-\zeta}}, \\
\widehat{p}_{bj,t}^{N_{bj,t}} &= \prod_{i=1}^{N_{bj,t}} \widehat{p}_{i,bj,t} \prod_{i=1}^{N_{bj,t}} \widehat{w}_{i,bj,t}^{\frac{1}{1-\zeta}} \prod_{i=1}^{N_{bj,t}} \widehat{s}_{i,bj,t}^{\frac{1}{\zeta-1}}, \\
\widehat{p}_{bj,t} &= \prod_{i=1}^{N_{bj,t}} \widehat{p}_{i,bj,t}^{\frac{1}{N_{bj,t}}} \left(\prod_{i=1}^{N_{bj,t}} \widehat{w}_{i,bj,t}^{\frac{1}{N_{bj,t}}} \right)^{\frac{1}{1-\zeta}} \left(\prod_{i=1}^{N_{bj,t}} \widehat{s}_{i,bj,t}^{\frac{1}{N_{bj,t}}} \right)^{\frac{1}{\zeta-1}}, \\
\widehat{p}_{bj,t} &= \widehat{\widetilde{p}}_{bj,t} \widehat{\widetilde{w}}_{bj,t}^{\frac{1}{1-\zeta}} \widehat{\widetilde{s}}_{bj,t}^{\frac{1}{\zeta-1}}, \\
\widehat{p}_{bj,t} &= \widehat{\widetilde{p}}_{bj,t} \widehat{\widetilde{s}}_{bj,t}^{\frac{1}{\zeta-1}}, \\
\widehat{p}_{bj,t} &= \frac{\widehat{\widetilde{p}}_{bj,t}}{\widehat{\widetilde{s}}_{bj,t}^{\frac{1}{1-\zeta}}},
\end{aligned}
$$

where $\widetilde{p}_{bj,t} \equiv \prod_{i=1}^{N_{bj,t}} \widetilde{p}_{i,bj,t}^{\frac{1}{N_{bj,t}}}$ is the geometric mean of unit values across product categories that buyer b sources from, and $\widetilde{s}_{bj,t} \equiv \prod_{i=1}^{N_{bj,t}} \widetilde{s}_{i,bj,t}^{\frac{1}{N_{bj,t}}}$ is the geometric mean of expenditure shares across products. Now, if we also introduce a time dimension into our estimating equation, express it in changes, and consider our expression for unit values, we have

$$PM_{i,bj,t}p_{bj,t}^{1-\zeta} = w_{i,bj,t}p_{i,bj,t}^{1-\zeta}PM_{bj,t},$$

$$\widehat{PM}_{i,bj,t}\widehat{p}_{bj,t}^{1-\zeta} = \widehat{w}_{i,bj,t}\widehat{p}_{i,bj,t}^{1-\zeta}\widehat{PM}_{bj,t},$$

$$\log\left(\frac{\widehat{PM}_{i,bj,t}}{\widehat{PM}_{bj,t}}\right) = (1-\zeta)\log\left(\frac{\widehat{p}_{i,bj,t}}{\widehat{p}_{bj,t}}\right) + \log\left(\widehat{w}_{i,bj,t}\right),$$

$$\log\left(\frac{\widehat{PM}_{i,bj,t}}{\widehat{PM}_{bj,t}}\right) = (1-\zeta)\log\left(\frac{\widehat{p}_{i,bj,t}}{\frac{\widehat{\widetilde{p}}_{bj,t}}{\widehat{\widetilde{s}}_{bj,t}^{\frac{1}{1-\zeta}}}}\right) + \log\left(\widehat{w}_{i,bj,t}\right),$$

$$\log\left(\frac{\widehat{PM}_{i,bj,t}}{\widehat{PM}_{bj,t}}\right) = (1-\zeta)\log\left(\frac{\widehat{p}_{i,bj,t}}{\widehat{\widetilde{p}}_{bj,t}}\right) + \log\left(\widehat{\widetilde{s}}_{bj,t}\right) + \log\left(\widehat{w}_{i,bj,t}\right).$$

D.2.2. Constructing price index $p_{i,bj,t}$

To estimate ζ, we need values for $p_{i,bj,t}$, which are not directly observed in the data since $p_{i,bj,t} \equiv \left(\sum_s \mu_{si,bj,t}p_{si,bj,t}^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}$, which is a function of ϵ and $\mu_{si,bj,t}$. For ϵ, we consider $\epsilon = \widehat{\epsilon}$, where $\widehat{\epsilon}$ is our estimated elasticity. For $\mu_{si,bj,t}$, we use the fact that the residuals when estimating ϵ are a function of these shocks. Recall that

$$\log\left(\frac{\widehat{PM}_{si,bj,t}}{\widehat{PM}_{i,bj,t}}\right) = (1-\epsilon)\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right) + X\beta + \phi_{si,bj,t},$$

where $\phi_{si,bj,t} = \log\left(\widehat{\mu}_{si,bj,t}\right) = \log\left(\frac{\mu_{si,bj,t}}{\mu_{si,bj,t-1}}\right) = \log\left(\mu_{si,bj,t}\right) - \log\left(\mu_{si,bj,t-1}\right)$ are the residuals of this estimating equation. By assumption, $\log\left(\mu_{si,bj,t}\right)$ are i.i.d and normally distributed shocks with mean μ and variance σ^2, so the mean and variance of $\log\left(\mu_{si,bj,t}\right) - \log\left(\mu_{si,bj,t-1}\right)$ is 0 and $2\sigma^2$, respectively. We now construct $p_{i,bj,t}$ by the following steps:

1. Estimate the 2SLS regression to obtain the estimate $\widehat{\epsilon}$;

2. Recover predicted values for the error term $\widehat{\phi}_{si,bj,t}$;

3. Calculate the empirical mean and variance of $\widehat{\phi}_{si,bj,t}$: $\left\{\widehat{\mu}_\phi, \widehat{\sigma}_\phi^2\right\}$;

4. Recover the values for mean and variance of $\log\left(\mu_{si,bj,t}\right)$, such that: (i) $\mu = \widehat{\mu}_\phi$ and $\sigma^2 = \frac{\widehat{\sigma}_\phi^2}{2}$;

5. Make a random draw for $\log\left(\mu_{si,bj,0}\right)$, which is drawn from a normal distribution with mean $\widehat{\mu}_\phi$ and variance $\frac{\widehat{\sigma}_\phi^2}{2}$;

6. For a given $\mu_{si,bj,0}$, recover $\mu_{si,bj,t}$ according to the following law of motion:

$$\begin{aligned}
\log\left(\frac{\mu_{si,bj,t}}{\mu_{si,bj,t-1}}\right) &= \widehat{\phi}_{si,bj,t}, \\
\frac{\mu_{si,bj,t}}{\mu_{si,bj,t-1}} &= \exp\left(\widehat{\phi}_{si,bj,t}\right), \\
\mu_{si,bj,t} &= \exp\left(\phi_{si,bj,t}\right)\mu_{si,bj,t-1};
\end{aligned}$$

7. We then construct unit values by

$$p_{i,bj,t} \equiv \left(\sum_s \mu_{si,bj,t} p_{si,bj,t}^{1-\widehat{\epsilon}}\right)^{\frac{1}{1-\widehat{\epsilon}}}$$

D.2.3. Constructing instruments

To obtain an exogenous shifter of relative unit values, which we use to obtain an unbiased estimate of ζ, we rely on the instruments we use to estimate ϵ. Consider the set of instruments $Z_{si,bj,t}$. Then, we consider the new set of instruments:

$$W_{i,bj,t} = \overline{Z}_{si,bj,t} = \frac{1}{N_{i,bj,t}} \sum_s Z_{si,bj,t}.$$

Consider the instrument that varies across both the color zone of the seller and the buyer (i.e. the share of districts in the *Red* zones within the route between the location of the seller and of the buyer). Then, the new instrument is the simple average of these shares across sellers. Intuitively, a higher share of districts in the *Red* zone should help predict a larger positive shock on unit values.

E SIMULATIONS USING QUANTITATIVE MODEL

E.1 Deriving expression for shock propagation through GDP

In this section, we discuss the details of the simulation using the quantitative model. First, recall the notations used in the paper. N is the number of firms, and I is the number of product categories. λ_k is the *Domar* weight of firm or sector k. θ_k is the elasticity of substitution corresponding to the k^{th} reproducible sector. Ω_{li} is the $(l,i)^{th}$element of the $(N+I+2)$ input output matrix Ω, which captures the direct reliance of l on i as a supplier . ψ_{li} is the $(l,i)^{th}$element of the $(N+I+2)$ Leontief inverse matrix $\psi \equiv (1-\Omega)^{-1}$, which captures the direct and indirect reliance of l on i as a supplier.

The aggregate change in GDP ($\Delta logy$) in response to changes in productivity of firm j ($\Delta logA_j$) up to a second order is given by the following:

$$\Delta logy = \sum_{j=1}^{N} \frac{\partial logy}{\partial logA_j}(\Delta logA_j) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1,i\neq j}^{N} \frac{\partial^2 logy}{\partial logA_i \partial logA_j}(\Delta logA_i)(\Delta logA_j) + \frac{1}{2}\sum_{i=1}^{N} \frac{\partial^2 logy}{\partial logA_i^2}(\Delta logA_i)^2. \tag{15}$$

Following Baqaee and Farhi (2019), after replacing second order terms, we obtain

$$\begin{aligned}
&= \sum_{j=1}^{N} \lambda_j (\Delta logA_j) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1,i\neq j}^{N} \left(\sum_{k=0}^{N} (\theta_k - 1)\lambda_k Cov_{\Omega(k)}(\psi_{(i)}, \psi_{(j)}) \right) (\Delta logA_i) \left(\Delta logA_j\right) \\
&+ \frac{1}{2}\sum_{i=1}^{N} \left(\sum_{k=0}^{N} (\theta_k - 1)\lambda_k Var_{\Omega(k)} \psi_{(i)} \right) (\Delta logA_i)^2 \\
&= \sum_{j=1}^{N} \lambda_j (\Delta logA_j) + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1,i\neq j}^{N} \left(\sum_{k=0}^{N} (\theta_k - 1)\lambda_k \left(\left(\sum_{l=1}^{N+F} \Omega_{kl}\psi_{li}\psi_{lj} \right) \right.\right. \\
&\left.\left. - \left(\sum_{l=1}^{N+F} \Omega_{kl}\psi_{li} \right) \left(\sum_{l=1}^{N+F} \Omega_{kl}\psi_{lj} \right) \right)\right) (\Delta logA_i) \left(\Delta logA_j\right) \\
&+ \frac{1}{2}\sum_{i=1}^{N} \left(\sum_{k=0}^{N} (\theta_k - 1)\lambda_k \left(\left(\sum_{l=1}^{N+F} \Omega_{kl}\psi_{li}\psi_{li} \right) - \left(\sum_{l=1}^{N+F} \Omega_{kl}\psi_{li} \right) \left(\sum_{l=1}^{N+F} \Omega_{kl}\psi_{li} \right) \right) \right) (\Delta logA_i)^2 \\
&= \sum_{j=1}^{N} \lambda_j (\Delta logA_j) + \frac{1}{2}B + \frac{1}{2}C.
\end{aligned} \tag{16}$$

We now write down the expressions for terms B and C in matrix form to evaluate second-order effects. In terms of notation, $J_{m,n}$ is a matrix of ones of size m by n, $\times$ is matrix multiplication, and $\cdot$ is element-by-element matrix multiplication.

Term B. This term primarily captures the second-order effects on GDP that operates through changes in firm $i's$ *Domar* weight in response to productivity shocks to firm j, where $j \in N, j \neq i$.

We construct this term in matrix form by sequentially deriving the following matrices:

$$
\begin{aligned}
&M = \psi \cdot (\Delta logA)^T, \\
&N = J_{(N+I+2,N+I+2)} \cdot \left(J_{(N+I+2,1)} \times \left(\psi \cdot (\Delta logA)^T\right)\right) - \left(\psi \cdot (\Delta logA)^T\right), \\
&Covar1 = \Omega \times (M \cdot N), \\
&Covar21 = \Omega \times M, \\
&Covar22 = \Omega \times N, \\
&Covar2 = Covar21 \cdot Covar22, \\
&B = \left((\theta - 1) \cdot \lambda\right) \times \left(Covar1 - Covar2\right).
\end{aligned}
$$

Term C. This term primarily captures the second-order effects on GDP that operates through changes in firm *i's Domar* weight in response to productivity shocks to firm *i* itself. In matrix form, this term is

$$
C = \left(\left((\theta-1) \cdot \lambda\right) \times \left(\Omega \times (\psi \cdot \psi) - (\Omega \times \psi) \cdot (\Omega \times \psi)\right)\right) \times \left(\Delta logA \cdot \Delta logA\right).
$$

GDP change in matrix form. In matrix form, we can rewrite Equation (15) as

$$
\begin{aligned}
\Delta logy = \lambda \times \Delta logA + .5\left((\theta-1) \cdot \lambda\right) \times \left(Covar1 - Covar2\right) + \\
.5\left(\left((\theta-1) \cdot \lambda\right) \times \left(\Omega \times (\psi \cdot \psi) - (\Omega \times \psi) \cdot (\Omega \times \psi)\right)\right) \times \left(\Delta logA \cdot \Delta logA\right) \quad (17)
\end{aligned}
$$

E.2 *Numerical implementation in Python*

Numerical implementations of our simulations are challenging due to the sheer size of the firm-to-firm trade network. We have data on 93,260 firms across 1293 product categories. This generates a 94,555 by 94,555 input-output matrix. The elements inside the input-output matrix are small as the fraction of a product's output to a single firm is small and, in turn, each product category sources from a large number of suppliers. To maintain calculations as precise as possible, we used $float64$ variable types within these matrices. Nevertheless, this also drastically increased the amount of computer memory required to hold matrices. For instance, the Leontief inverse matrix required more than $66GB$ of storage/memory size.

These large matrices require many steps to perform matrix multiplication operations on them. In computer science, matrix multiplication is one of the most demanding operations in terms of computing resources. We break down these operations by leveraging state-of-the-art computing techniques in big data. These techniques provide us with scalability when applying them to arbitrarily large input-output matrices. As detailed firm-to-firm transaction data are becoming more widely available, these techniques are promising to advance the literature on quantifying the propagation of shocks through firm networks. We now briefly describe these techniques.

First, we fit datasets larger than RAM using Dask, which is a Python library with multi-core, distributed, and parallel execution on larger-than-memory datasets.[29] We use Dask's distributed capabilities to parallelize our calculations when computing second-order effects which require few matrix multiplication operations on large 94,555 by 94,555 matrices.

Second, we use a computer powered by multiple GPUs. GPUs are essential for performing a large number of matrix multiplications. For example, computing 10 columns of the Leontief inverse matrix, which is only around 0.0001% of columns we need to compute, takes about 4 days on a powerful server with multiple CPUs, 500 GB of RAM, and 16 cores. Computing the entire Leontief inverse on a server powered with 4 GPUs takes about 1 hour.

Third, we use the properties of sparse matrices to define matrix multiplications that ignore large contiguous sub-matrices full of zeros, which is a typical feature of input-output matrices.

Fourth, we developed a custom matrix multiplication function to overcome the limitation of the relatively small memory size of GPUs. The custom matrix multiplication function splits the matrix into sub-matrices of full columns (typically in the order of a few 1000's of columns), it multiplies the sparse input-output matrix by each sub-matrix, and it concatenates all result chunks to formulate the final result.

[29] https://tutorial.dask.org/00_overview.html