Production Networks and Firm-level Elasticities of Substitution*

Brian Cevallos Fujiy† Devaki Ghose‡ Gaurav Khanna §

Preliminary and Incomplete

Abstract

We use geographic and temporal variation from the Covid-19 lockdowns in India to quantify the fall in trade and estimate elasticities of substitution at the firm-level. Using new real-time administrative tax data on firm-to-firm transactions, we provide one of the first estimates of elasticities of substitution across inputs supplied by suppliers within the same HS-2 industry. This estimate is particularly relevant for the transmission of supply shocks. Not having access to suppliers from strict lockdown zones forced firms to look for alternative suppliers. If suppliers in the lockdown and non lockdown zones are complements rather than substitutes in production, this shock can amplify by further transmitting downstream through the supply chain. We find that even at this very granular supplier level, inputs are highly complementary, with an estimated elasticity 0.38. Causally estimating these micro-level elasticities of substitution at the firm level allows us to understand how shocks propagate through supply chains, and how these propagations affect aggregate GDP.

Keywords: production networks, elasticity of substitution
JEL Codes: F41, F44

*We thank Eli Mogel for excellent research assistance, and Prof. Sushil Khanna for guidance on the data and details on the analysis. We thank seminar participants at the Bank of Italy, and Ludovic Panon for insightful comments. The views expressed in this paper do not represent the views of World Bank or its partner organizations, and solely represent the personal views of the authors.
†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 find suppliers is key for the resilience of supply chains. This is particularly relevant for the transmission of supply shocks. For example, if it is difficult for firms to substitute across suppliers after a negative supply shock, this shock could amplify by further transmitting downstream through the supply chain. In many developing countries during the Covid-19 pandemic lockdown policies were accompanied with large GDP declines.[1] For instance, India reported a -7.3% growth rate for the 2020/21 financial year, one of the largest contractions worldwide and since its independence.[2] In this paper, we quantify the importance of firm-level elasticities of substitution across suppliers of intermediate inputs to explain large fluctuations in GDP. We provide new estimation strategies and estimates for these elasticities by leveraging regional variation in trade costs induced by the Indian government's massive lockdown policy to address endogeneity concerns. We hypothesize that this elasticity could be key to partly explaining the large decline of the Indian economy during the Covid-19 pandemic.

We pose two main research questions. First, are inputs across suppliers in closely related industries complements or substitutes? The answer to this question is important since this determines how shocks propagate throughout supply chains. If inputs are substitutable, we would expect shocks to propagate less across firm networks. However, if inputs are complements, the effects of negative shocks can easily propagate through buyer-supplier networks. Second, how do these estimated elastícties affect firm-level sales, and hence, ultimately GDP, by propagating and amplifying shocks through firm-level input-output linkages? The current version of the draft mainly focuses on the first question, that is, on consistently estimating the firm level elasticity of substitution across suppliers.

Two unique features of our setting allows us to answer these questions in a credible manner. First, India had a distinct lockdown policy, whereby the roughly 600 districts were classified into three difference zones with varying degrees of restrictions. This allows researchers to derive variation in the ability to trade and transport goods over this period. Second, we are able to obtain new granular and high-frequency administrative data on the universe of firm-to-firm transactions for a region in India. These data, while not used before, allow us to estimate elasticities at the firm (rather than industry) level, and across different suppliers for a firm.

We find that inputs within the same HS-2 industry but across different suppliers are highly complementary, with an estimated value of 0.4. In various robustness checks employing different combinations of fixed effects and using different time periods for identification, we find that the estimated elasticities lie within a range of 0.37 to 0.41. These estimates are only slightly higher than Atalay (2017) who estimates this to be around 0.2, but at the level of the industry,

[1] -3.3% and -2.2% growth rates during the 2020/21 financial year for emerging market and developing countries, respectively.

[2] https://www.economicsobservatory.com/how-has-Covid-19-affected-indias-economy

and we show that even within the same HS-2 industry, inputs across firms are highly complementary. This suggests that even at the very micro-level, firm-specific negative shocks are capable of contributing to fluctuations in GDP.

Estimating elasticities of substitution across different suppliers has been especially challenging for two reasons. First, as discussed by Taschereau-Dumouchel (2020) and Baqaee and Farhi (2019), the literature so far provides very little guidance about estimates of the firm-level elasticity of substitution between suppliers of intermediate inputs. While estimates of elasticities of substitution across intermediate goods across industries have recently been estimated (Atalay, 2017), such estimates do not exist for suppliers within the same industry. The reason for this gap in the literature lies in the difficulty of finding detailed data on firm-to-firm input transactions, as well as finding an exogenous source of variation in firm level prices that allows to properly estimate these elasticities. We provide estimates of firm-level elasticities of substitution across suppliers by leveraging the nation-wide sudden and unprecedented lockdown imposed by the Indian government starting in March 2020. Depending on the severity of Covid-19 cases, districts were categorized into *Green* (mild lockdown), *Orange* (medium lockdown) and *Red* (severe lockdown). The fact that the lockdowns were sudden and unexpected,[3] were implemented independent of economic fundamentals across districts, and induced strong variation in transactions between firms across India, help us exploit this shock to estimate the firm-level elasticities.

The second challenge is related to the fact that the Covid-19 is not just a supply shock. As pointed out by Baqaee and Farhi (2020), the outbreak of the pandemic is a combination of exogenous shocks to the quantities of factors supplied, the productivity of producers, and the composition of final demand by consumers across sectors. To estimate the elasticity of substitution across the suppliers of inputs for a particular product produced by a firm, we leverage variations in input prices predicted by the sudden restrictions in economic activities due to lockdowns in the districts where these suppliers are located. In addition, we also leverage the variations in trade costs arising from restrictions in economic activities in districts through which the goods need to pass from the seller to the buyer. To further isolate supply shocks from other shocks, we control for various other factors, such as constructing firm-level exposure to foreign shocks transmitted through trade following Hummels et al. (2014), and the caseload and severity of Covid-19 cases.

This paper has three main sections. First, we present-reduced form evidence on the impact of negative supply shock on key firm-level variables such as unit values (prices) and number of transactions (quantity). We leverage the Indian government's sudden lockdown measure that affected firm-to-firm trade across districts, depending on whether firms fall in the red zone (strict lockdown), orange zone (moderate lockdown) or green zone (mostly no lockdown). We

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

find that the prices of intermediate inputs rose during the lock-down, especially if either the buyer or the seller was located in the orange or red zone. In districts where the seller is located in a strict lockdown zone (orange or red), the number of transactions fell drastically, compared to either the case where the buyer is located in a lockdown zone or the case where both are located in green zones.

Second, we modify a standard multi-sector firm level model of input-output linkages by augmenting the production functions with substitution across suppliers within the same industry. From the model we generate analytical expressions that relate the relative values of quantities purchased of the same HS-2 product from different suppliers to the equilibrium relative prices. That is, within each HS-2 product category, we quantify how substitutable are the inputs from the different suppliers. This helps us estimate the elasticities of substitution at the very short-run. The literature has consistently showed that this short-run elasticity is near zero while the long-run elasticity is non-zero. We hypothesize that while this elasticity could be near zero when we look at broad product categories, it could be much higher if we look at firm-level elasticities. We find that this elasticity is close to 0.40. Thus, following Baqaee and Farhi (2020), after considering second-order effects, negative firm-level shocks get amplified in the aggregate by propagating through firm-to-firm linkages while positive shocks get dampened. We further explore whether these elasticities differ across industries and find that in a few handful of industries, suppliers within the same industries are actually substitutes rather than complements. This shows that we should be mindful of heterogeneity across industries in understanding how shocks propagate through supply chains.

In future versions, we plan to include a third section where we use the estimated elasticities to analyze how input complementarities at the firm level affect aggregate economic outcomes, and thus, how important these complementarities are in explaining the GDP decline in India during the Covid-19 pandemic. Armed with these new elasticity estimates, we are also able to analyze how shocks to even single firms in the production network can affect other firms, depending on the affected firms' size and its direct and indirect linkages throughout the economy. We will then compare how these magnitudes differ if we were to ignore these firm level elasticities of substitution across suppliers.

2 Related Literature

Our paper connects with three broad strands of literature. First, we speak to the literature on shock propagation through supply-chains. Understanding how shocks propagate from one firm to another firm has received relatively less attention in the literature, primarily due to lack of data, and particularly due to the lack of identifying variation in firm-specific shocks (Barrot and Sauvagnat, 2018). Relative to their work, which mostly looks at disruptions due to natural disasters on firms' customers and on customers' other suppliers, we can directly trace out how shocks propagate through firm-to-firm supply chain relationships as we can observe firm-to-firm input trade.

Relying on firms' location, Carvalho et al. (2021) exploit the heterogeneous exposure of Japanese firms to the earthquake to obtain measures of firm-level disturbances. They combine this information with extensive micro-data on inter-firm transactions to trace and quantify the extent of shock propagation along supply chains. In addition to the binary measure of inter-firm supplier-customer relations that they observe, we crucially observe the intensive margin or the number of transactions in a single relationship. We directly use information on firm-to-firm input sales to measure elasticities of substitution at the firm level. This relates our work to Peter et al. (2020) and Boehm et al. (2019) who also estimate elasticities of substitution between inputs in the long and short-run, respectively. While their work provide estimates of elasticities of substitutions between material inputs across different industries, we provide this estimate at a substantially more micro level – the elasticity of substitution between inputs supplied by different firms within the same HS-2 industry.

Second, this paper is closely related to empirical research on trade collapse during large negative shocks. Using firm-level Belgian data, Behrens et al. (2013) find a comparable collapse of domestic and cross-border operations due to the financial crisis. Giovanni and Levchenko (2009) examine variation in US exports and imports across 6-digit industries and find that industries experiencing larger reductions in domestic output also had a larger fall in trade. Using monthly firm-level exports from France, Bricongne et al. (2012) find a dominant role for the intensive margin fall in trade. Indeed, as Baldwin and Tomiura (2020) argue "2020 will show a trade collapse that is far larger since the 'Covid concussion' is both a demand shock and a supply shock while the 2008-09 collapse was driven mostly by a demand shock. In today's Covid-19 crisis, we have all the makings of the 2008-2009 demand side shock, but on top of that we have massive, supply-side shocks across most sectors of most major economies."[4] Along the lines of this observation, we think that it is important to understand how the current crisis will affect both internal and external trade. To our knowledge, this is the first study to analyze how the Covid-19 crisis, along with the government mandated lock-downs, affected internal trade using detailed firm-to-firm transactions.

This relates our paper to the literature that studies the transmission of shocks through GVCs during the Covid pandemic mainly by looking at disruptions to firm level imports and exports or aggregate production due to the crisis (Bonadio et al., 2021; Baqaee and Farhi, 2020; Cakmakli et al., 2021; Demir and Javorcik, 2020; Gerschel et al., 2020; Heise et al., 2020; Lafrogne-Roussier et al., 2021). We complement this literature by analyzing the impact of the Covid crisis on domestic firm-to-firm trade. We document how domestic firm-to-firm transactions were affected following the imposition of lockdowns due to Covid in a large developing

[4] VoxEU Article: https://voxeu.org/article/greater-trade-collapse-2020

country, India, and then use the lockdown events to study whether suppliers of a firm within the same industry are substitutes or complements. The key policy motivation behind this project is the observation that policy makers worldwide are interested in better understanding the trade offs between strict economic lock-downs that prevent the spread of the virus but can affect GDP growth through complex buyer seller networks and more lenient measures that increase production and trade but can lead to potentially wider spread of the virus. More importantly, even beyond the immediate Covid crisis, the estimates of how substitutable or complementary suppliers are within a given industry, will help policy makers quantify the economy-wide effects of any disruptive events (e.g natural disasters or policies such as lockdown) on trade and production.

Finally, using previously unavailable data on firm-to-firm transactions from a large Indian state we document new facts on firm-to-firm trade and production networks. This relates our work to the small but burgeoning literature that use detailed firm-to-firm domestic transaction data to study various features of the production network (Panigrahi, 2021; Demir et al., 2021; Dhyne et al., 2021; Alfaro-Urena et al., 2020).

3 Data and Context

This section provides a general overview of our data and context.

Firm-to-firm trade. Our primary data source is daily establishment level transactions.[5] This data is provided by the tax authority of a large Indian state with a fairly diversified production structure, roughly 50% urbanization rates, and high levels of population density. Comparing this context to other contexts with firm-to-firm transaction data, we observe that the state has 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. Each transaction reports a unique tax code identifier for both the selling and the buying establishments. Each transaction reports all the items contained within the transaction, the value of the whole transaction, the value of the items being traded by 8-dgit HSN code, quantity of each item, its unit, and the mode of transportation.

Each transaction also reports the pincode (zip code) location of both the selling and buying establishments, which we use as a key to merge with other district-level data. By law, any person dealing with the supply of goods and services whose transaction value exceeds 50,000 Rs (700 USD) will have to generate eway-bills. Transactions that have values lower than 700

[5]While we use the term 'firm' in most parts of the paper, these data are actually at the more granular establishment level, and we can identify the parent firms for each establishment as well.

USD can also be registered but it is not mandatory. Our data is generated from these eway bills. This implies that our network is certainly representative of relatively larger firms, but this threshold is sufficiently low such that we are confident we are capturing small firms as well.

The data does not contain information about prices, but it does report value and quantity of traded items, so we can construct unit values. To do this, we aggregate values and quantities at the 2-digit HSN/month/transaction level, and then construct implied unit values. We can then collapse the data at the 2-digit HSN/month level to construct average unit values, the number of transactions between each seller and buyer pair, and total value of the goods transacted. This is the foundation of our firm-to-firm dataset that we use in the analysis. Additionally, we can aggregate this data to the seller level, which we use in our reduced-form section.

Each establishment is located within a district, so treated firms are located within a *Red*, *Orange*, or *Green* district in April or May 2020 according to Indian lockdown policies, the details of which are given below.

Lockdowns. On March 25th 2020, India imposed strict lockdown policies nation-wide. These policies were unexpected and their duration was indeterminate. The lockdown was implemented nation-wide at the district level, where each district was classified between *Red*, *Orange*, and *Green* according to the severity of Covid cases in each district, and thus, the severity of lockdown measures. In Figure 1 there is a map showing the distribution of lockdowns across India. Districts in the red zone saw the strictest lockdown measures, with rickshaws, taxis and cabs, public transport, barber shops, spas, and salons remaining shut. E-commerce was allowed for essential services. Orange and green zone districts saw fewer restrictions. In addition to the activities allowed in red zones, 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.[6]

Physical and Cultural Distance. The measures of geographic distance between districts are obtained from Kone et al. (2018) who calculate the length of the shortest distance between district centers. The measure of cultural distance used in the data is essentially a measure of linguistic distance between Indian districts. This is also obtained from Kone et al. (2018) who construct linguistic distances between any two districts (i, j) following the commonly used ethno-linguistic fractionalization (EFL) index (Mira, 1964). This index measures the probability of two randomly chosen individuals from different districts speaking the same language. More details on this measure is given in Kone et al. (2018).

[6]https://economictimes.indiatimes.com/news/politics-and-nation/lockdown-3-0-guidelines-for-red-zone/activities-prohibited/slideshow/75503925.cms

Controls. We control for different firm and district level time varying variables such as data on monthly number of cases, deaths, and recoveries from Covid-19 for all India at the district level from `www.Covidindia.org`.[7] For each firm, we construct two variables that measure the firm's exposure to global demand and supply shocks that vary at the HS-2 product and country level, following Hummels et al. (2014). The construction of these exposure variables are described in detail in online data appendix A.

Summary statistics : We present some key summary statistics from the administrative trade data in table 1. Panels A and B report the unique numbers of sellers, buyers, total sales (in million rupees), and total number of transactions separately in months January-March, April-June, and July-September, for years 2019 and 2020. The most noticeable pattern from the data is the large drop in all variables in 2020 in comparison to 2019, particularly during the April-June period, which coincided with the lockdown policies.

Compared to April-June of 2019, the total number of sellers and buyers fell by 47.5% and 39.4% respectively in the corresponding months of 2020. Moreover, the total value of sales and the number of transactions both 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.

4 Reduced form evidence

In this section we describe our empirical strategy, and then provide evidence showing the role of lockdown policies in India on key outcome variables for firm-to-firm trade. We show that the sudden lockdown policies in India led to a rise in unit values, and a fall in the monthly number of transactions between firms. We leverage quasi-exogenous variation from Indian lockdown policies in April-May 2020 to estimate the effect of lockdown policies on prices and quantities of intermediate inputs between firms.[8]

4.1 Empirical specification

Our main reduced form specifications employ difference-in-difference specifications where we compare the unit values and the number of transactions both at the seller-buyer pair level and at the seller-level across red, green, and orange districts, before and after the lockdown. In

[7]The exact link to the data is `https://docs.google.com/spreadsheets/d/1lgaEhEPfXiLr-88QgtBrEoE-m-lPIpKuIZS7E80EBlY/edit#gid=1493892497`

[8]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).

our analysis at the seller-level, the control group are sellers located in *Green* districts and the base month is February 2020, two months before the enforcement of lockdown policies. At the seller/buyer level, the control group are sellers and buyers located in *Green* districts and the base month is February 2020.

4.1.1. Seller-level regressions

We estimate the following specification:

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

where $Y_{si,t}$ are unit values or the number of transactions in logs for seller s in HS-2 sector i in month t, ι_i are 2-digit HSN fixed effects, $\iota_{o_{(s)}}$ are district fixed effects (i.e. fixed effects based on the district o where seller s resides), ι_t are month fixed effects, X are controls that include number of Covid cases, deaths, and recoveries, and exposure to international demand and supply shocks as discussed in Appendix A. It is important to 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 one is a dummy variable that equals 1 if seller s located in district $o_{(s)}$ experienced a severe lockdown, 0 otherwise. The second one equals 1 if seller s located in district $o_{(s)}$ experienced a mid-level lockdown, 0 otherwise. The excluded category are $Green_o$ districts, where mild lockdown was imposed. The estimators of interest are β_t and γ_t. Our base time category is February 2020 which is just before lockdowns began.

4.1.2. Seller-buyer level regressions

At the seller-buyer level we estimate the following specification:

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

where $Y_{si,b,t}$ are unit values or number of transactions in logs between seller s in HS-2 sector i and a buyer b in month t. $\delta_{o_{(s)}}$, $\delta_{d_{(b)}}$, δ_i, and δ_t are origin, destination, sector, and month fixed effects. $dist_{od}$ is a vector of cultural and geographic distance variables, and X are controls that include number of Covid-19 cases, deaths, recoveries and exposures to international demand and supply shocks. The first term of the right-hand side requires further explanation since it contains our estimators of interest. $(x,z) \in \Omega$ is a duple that contains the color x 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. $\gamma_{o_{(s)}}^x$ and

$\gamma^{z}_{d_{(b)}}$ are thus dummy variables that equal 1 when seller s is located in district o located in lockdown zone x, and when buyer b is located in district d located in lockdown zone z, respectively. The estimators of interest are β^{xz}_{t}. Our base time category is February 2020 which is just before lockdowns began.

4.2 Results

In this section we present our main results from the specifications we laid out in the previous section.

Fact 1: Unit values rose during Covid-19 lockdowns. Unit values unambiguously rose during the Covid-19 lockdown in India. We can see this either at the seller level, or at the seller-buyer level. In the first panel of figure 3a which plots the coefficients from the seller level regression in equation 1, we can see that, in comparison to sellers in green districts, sellers in orange and red districts witnessed an increase in unit values of around 20pp during the lockdown month of April. Figure 4, which plots the coefficients from the seller-buyer level regression in equation 2, shows that if either the buyer or the seller is located in an orange or red zone, the unit values associated with their transactions rose, compared to a situation where both buyers and sellers are located in green zones during the lock down period. In both the figures, we find no evidence of pre-trends, meaning that there were no significant differences in the unit values or in the number of transactions between red, orange, and green districts before the lockdown.

Fact 2: The rise in unit values increases with the severity of the lockdown. This can be observed very clearly when studying the results at the seller-buyer level. In the first row of figure 4 we plot the coefficients from regression 2 where the seller is in red zone, and the buyer is in red, orange, and green zones respectively. In the second row of figure 4, we plot the coefficients from regression 2 where the seller is in orange zone, and the buyer is in red, orange, and green zones respectively. In the third row, we plot the coefficients from regression 2 where the seller is in green zone, and the buyer is in red and orange zones respectively. We notice that, the rise in unit values increases in the severity of the lockdown – ranging from 25 pp to 55 pp – depending on whether we compare a supplier-buyer pair in an orange-red or a orange-green district pair, to a supplier-buyer pair in the green-green district. All other cases lie in between. Notice that since the number of observations (which corresponds to the number of unique monthly transactions at the seller-buyer-HS2 level) were too small for trade between suppliers and buyers where both are located in red districts, the corresponding coefficient for transactions between buyers and sellers in red districts is less precisely estimated. We see similar results in figure 3a, where the unit values rose more when the seller was in red district compared to when the seller was in orange district, but the results are less stark since we have aggregated the results to the seller level. Collectively, the figures provide suggestive evidence that supply shocks brought about by lockdowns increases prices in varying degrees depending on the degree of severity of the lockdowns.

Fact 3: Number of transactions went down during Covid lockdowns. The number of transactions plummeted during the Covid-19 lockdown in India. We can see this either at the seller level, or at the seller-buyer level. In the second panel of figure 3b which plots the coefficients from the seller level regression, we can see that, in comparison to sellers in green districts, sellers in orange and red districts exhibited a large decrease in the number of transactions, particularly in red districts where they dropped around 20pp. In figure 5, which plots the coefficients from regression 2, we see that the number of transactions dropped only in instances where the seller was in an orange or red district. This also provides suggestive evidence that supply-side shocks induced by the lockdown affected the number of transactions between firms.

Fact 4: The decline in the number of transactions increases with the severity of the lockdown. This can be observed very clearly when studying the results at the seller-buyer level regression. As observed in fact 3, the decline in the number of transactions occur mainly when the seller is in an orange or red district. Then, if we observe figures 5a, 5b, 5c, we see that conditional on the seller being in a red district, the magnitude of the coefficients increase with the severity of the lockdown from the buyer side. From figure 3b, which plots the coefficients from the seller level regression, we can see that the fall in transactions was much more severe when the seller was in the red zone compared to when the seller was in the orange zone.

5 Quantitative Framework

In the previous section we provide evidence that lockdown-induced supply shocks led to a rise in unit values and a fall in transactions, especially when the seller was in red or orange districts. Based on this findings, we adapt the general nested CES structure from Baqaee and Farhi (2019) to reflect the possibility that suppliers within the same industry could be substitutes or complements, derive the estimating equations to estimate the elasticities in this framework, and discuss how we plan to use the model to conduct simulations in future versions. We follow the framework in Baqaee and Farhi (2019) to write out a nested CES economy in standard form.[9]

There are N firms producing N goods using the production function:

[9] We do not rely on models featuring market power such as (Edmond et al., 2018; Alviarez et al., 2021) since the evidence from the data suggests that the market structure in this Indian state is more towards perfect competition. Unconditional on the HSN industry, the HHI is of 0.004, which indicates a highly competitive market. When we calculate HHIs by HSN sections, the median industry by HHI exhibits a value of the index of 0.013, which still implies a high level of competition.

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

where n denotes the firm and j denotes the industry the firm belongs to. l_n denotes the labor used by firm n, x_{nj} is the composite intermediate input used by firm n in industry j, α is the elasticity of substitution between labor and the composite material input and w_{nl} denotes the intensity of labor in production. The composite material input in turn consists of inputs from the I different industries in the economy, and is given by:

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

ζ denotes the elasticity of substitution between inputs from different industries. $x_{i,nj}$, defined below in equation 5, are intermediate inputs from industry i going to firm n in sector j. We do not distinguish between foreign inputs from different firms as we do not have any data on them.

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

$x_{mi,nj}$ denotes intermediate inputs from firm m in industry i going to firm n in industry j. N_m denotes the number of firms in industry m. $\sum_k N_m = N$ denotes the total number of firms in the economy. ζ denotes the elasticity of substitution between different inputs from different HS-2 digit industries, and ϵ denotes the elasticity of substitution across firms within the same industry. In the next section we will estimate this latter elasticity. The above production functions work for reproducible factors. For non-reproducible factors, in our case, labor, the production function is an endowment: $Y_f = 1$.

Industry 0 represents the final consumption of the household and is given below:

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

where $\sum_i w_{oi} = 1$ and σ is the elasticity of substitution in consumption. To write the economy in standard form such as in Baqaee and Farhi (2020), we define a new input output matrix $\hat{\Omega}$ which has dimension $2 + N + I$, where the first dimension represents the household's consumption aggregator, the next dimension correspond to factors, here only labor, the next N dimensions are the N firms that supply inputs to the CES aggregates and the next I dimensions are the CES aggregates of intermediate inputs of these firms that directly go into the firm's production function. Let us denote the vector of elasticities by $\hat{\theta}$, where $\hat{\theta} = (\alpha, \zeta, \epsilon, \sigma)$.

Formally, a nested-CES economy in standard form is defined by $(\hat{\Omega}, \hat{\theta})$. What distinguishes factors from goods is that factors cannot be produced. The $(2+N+I)\times(2+N+I)$ input–output matrix $\hat{\Omega}$ is the matrix whose (i,j) element is equal to the steady-state 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 share of the total revenue of firm i, where, note that, every supplier is a CES aggregate. The Leontief inverse is $\psi=(1-\Omega)^{-1}$. Intuitively, the $(i,j)th$ element of ψ of the Leontief inverse is a measure of i's total reliance on j as a supplier. 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, introduced in Baqaee and Farhi (2019) is given by:

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

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 detailed derivation of this, see the appendix of Baqaee and Farhi (2019). To get an intuition of how firm-level shocks can propagate through supply chains, consider a specific example: firm j, located in the red zone, suffers a negative productivity shock, given by $d\log A_j<0$. This negative productivity shock to a specific producer j changes the vector of prices of the different producers in the economy depending on their direct and indirect exposures to the shock, summarized by column j of the Leontief inverse ψ. How that changes the sales of producer i arise from the substitution by the different producers k in the economy and are captured by the different terms in the sum on the right-hand side.

Now consider a given producer k. If $\theta_k<1$, producer k increases its expenditure share on inputs from firms whose prices increase more, that is, inputs that are more exposed to the shock to j, as measured by $\psi_{(j)}$. If these same firms are also directly and indirectly exposed more to firm i, as measured by $\psi_{(i)}$, the relative demand expenditure for inputs from i also increases. The overall effect is stronger, the higher is the covariance $Cov_{\Omega_k}(\psi_{(i)},\psi_{(j)})$, the larger is the size of producer k as measured by λ_k, and the further away from 1 is the elasticity of substitution θ_k. We would sum up the effects across all producers k in the economy to find the overall effect of the productivity shock of firm j on firm i, taking into account all possible input output linkages in the economy. Thus, the second order effect of the shock on the aggregate GDP depends on the firm level micro-economic elasticity of substitution. High complementarity between inputs supplied by different firms amplify the negative effects of a supply shock to a specific producer in the economy, and the magnitude of that amplification depends on the size distribution of firms in the economy and the economy wide linkages of that firm with other firms, as given by the Leontieff inverses.

5.1 Deriving the estimating equation for the elasticity of substitution across firms

Using the model outlined above, in this section we derive the elasticity of substitution at the firm level. The only notational change from the previous section that we introduce here is that a firm n can be either a buyer (b) or a seller (s). Consider a discrete set of firms F and a discrete set of sectors J. A seller is denoted by $s \in F$ and a buyer by $b \in F$. A firm b in sector $j \in F$ 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 (3) , (4), and (5). ϵ is the elasticity of substitution across different suppliers within the same industry. This is the key elasticity we want to estimate. Note that the results of this estimation procedure holds with any CES production function with an arbitrary number of nests, as long as the lowest nest consists of suppliers within the same HS-2 industry.

Details about the optimization problem are in Appendix C.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 $PM_{i,bj} \equiv \sum_s PM_{si,bj}$, and $\log\left(\mu_{si,bj}\right)$ is the error term. This is our main estimating equation for the firm-level elasticity of substitution parameter ϵ which we take to the data, as will be described in detail in section 6.

6 Estimation and Quantification

In this section, we discuss how we estimate the unknown parameters and quantify the model to understand how firm-level shocks can propagate through firm GVC networks. The vector of unknown parameters is given by $\hat{\theta} = (\alpha, \zeta, \epsilon, \delta)$. We set the elasticity of substitution between different consumption varieties $\sigma = 4$, the elasticity of substitution between labor and the composite intermediate input $\alpha = 0.5$, the elasticity of substitution between different varieties $\zeta = .2$, drawing on Feenstra et al. (2018), Atalay (2017) and Baqaee and Farhi (2020). There are no known values of the firm-level elasticity of substitution in the literature and we estimate this parameter ϵ using equation 9 first by OLS and then by using the Indian lockdown policies as instruments to address potential endogeneity concerns for our estimators.

6.1 Addressing unobservable productivity shocks

In Equation (9), our main estimating equation for the firm-level elasticity of substitution, note that $p_{i,bj,t}$ includes the unobserved productivity shock $\mu_{si,bj,t}$. Following Redding and Weinstein (2020), we can arrive to the following expression which we directly take to the data:

$$\log\left(\frac{\widehat{PM}_{si,bj,t}}{\widehat{PM}_{i,bj,t}}\right) = (1-\epsilon)\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\tilde{p}}_{i,bj,t}}\right) + \omega_{d(b),t} + \omega_{o(s)} + X\beta + \epsilon_{si,bj,t}. \tag{10}$$

For more details about how we arrive to this expression, see Appendix C. $\omega_{d(b),t}$ are destination district-by-month fixed effect, and ω_{o_s} are origin district fixed effects. X are controls, including exposure to foreign demand and supply shocks, the number and severity of Covid-19 cases, and geographic and cultural distances. First, notice that we introduced a time dimension, so now variables exhibit a t subscript. We then use the notation $\hat{x}_t \equiv \frac{x_t}{x_{t-1}}$ to express variables in changes with respect to the previous month. Under the assumption that the geometric mean of productivity shocks across suppliers of buyer b are constant, we can then rewrite the main right-hand side variable as $\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\tilde{p}}_{i,bj,t}}\right)$. In this variable, notice that the denominator is now $\tilde{p}_{si,bj,t}$ instead of $p_{si,bj,t}$, where

$$\tilde{p}_{i,bj,t} = \prod_s p_{si,bj,t}^{\frac{1}{N_{i,bj,t}}}$$

is a geometric mean across suppliers. We can then run OLS estimators of the elasticity of substitution across suppliers. The main advantage of this setup is that we can construct this variable straight from the data. We also include controls X described in Section 3 and fixed effects at the destination state/month and origin state level. Standard errors are clustered at both the origin state level, and origin-destination state pair level. The logic behind clustering at the state level is that most policies, even those announced by the center, are implemented at the state level.

6.2 Addressing endogeneity concerns

OLS estimates of ϵ are biased if unobserved changes in production technologies drive changes in prices and expenditure shares. Measurement error in input prices, proxied by unit values, can also create attenuation bias. 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 can also bias our estimation. For example, if Covid-19 induced demand shocks lead to contractions in buyers income and at the same time supply-shocks lead to contractions in the sellers supply, the demand curve will look flatter (estimated ϵ higher) compared to the unbiased value of ϵ.

Our estimation strategy therefore involves using sudden demarcations of lockdown zones that restrict economic activities in certain Indian districts as an instrumental variable when estimating this equation in 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 supplies. The idea is that after controlling for the lockdown zones the buyer is located in, exposures to international demand and supply shocks, the number and severity of regional Covid-19 cases, the 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 lock-downs provide another source of exogenous variation to estimate the firm-level elasticity of substitution.

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

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

In logs, this expression is

$$\log\left(p_{si,bj,t}\right) = \log\left(\tau_{sb,t}\right) + \log\left(p_{si,t}\right).$$

Here we can see the types 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. These help us obtain unbiased estimates of the elasticity ϵ. Second, exogenous shifters at the seller-buyer level, for example, changes in transportation costs $\tau_{sb,t}$ driven by the lockdown zones of the districts the goods pass through, could also induce the needed variation. Empirically, we will use these instruments separately and jointly. We now describe each of these instruments.

Seller-level instruments. We need supply-side shifters to obtain unbiased elasticities of substitution. In that sense, shocks induced by the Covid lockdown policies that only impact sellers would provide that variation. In equation 11 below we formalize this intuition.

$$log(\widehat{p}_{si,bj,t}) = \beta^R Red_{o_{(s)}} Lock_t + \beta^O Orange_{o_{(s)}} Lock_t + \omega_{d_{(b)},t} + \omega_{o_{(s)}} + X\beta + \epsilon^{\nu}_{si,bj,t}. \qquad (11)$$

$Lock_t$ is a dummy variable that equals 1 for the months of April and May of 2020, which are the months when the lockdown policies were implemented, 0 otherwise. $Red_{o_{(s)}}$ and $Orange_{o_{(s)}}$ are dummy variables that equal 1 whenever seller s was located in *Red* or *Orange* districts, respectively. $\omega_{d_{(b)},t}$ are destination district-by-month fixed effects, $\omega_{o_{(s)}}$ are origin district fixed effects, and X are the same set of controls as in Equation 10.

Seller/Buyer-level instruments. The transportation of supplies from the location of the supplier to the buyer implies going through different districts, each of which are 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. Notice we allow trade cost to change over time such that we can leverage the Covid lockdown policy. In particular, we assume

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

That is, Covid lockdown is an exogenous shifter that only influences travel time between locations of seller s and buyer b. If we consider this to be the only shifter, and after considering this functional form for trade costs into the expression of prices and log-differences, we obtain

$$\log(\widehat{p}_{si,bj,t}) = \sigma \log(\widehat{traveltime}_{sb,t}).$$

In particular, we consider

$$log(\widehat{p}_{si,bj,t}) = \beta^R Red_{o_{(s)}d_{(b)}} Lock_t + \beta^O Orange_{o_{(s)}d_{(b)}} Lock_t + \omega_{d_{(b)},t} + \omega_{o_{(s)}} + X\beta + \epsilon^{\nu}_{si,bj,t}. \tag{12}$$

Details on how we obtain this functional form are contained in Appendix C.3. $Lock_t$ is a dummy variable that equals 1 for the months of April and May of 2020, which are the months when the lockdown policies were implemented, and 0 otherwise. $Red_{o_{(s)}d_{(b)}}$ and $Orange_{o_{(s)}d_{(b)}}$ are the share of 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 A. $\omega_{d_{(b)},t}$ are destination district-by-month fixed effects, $\omega_{o_{(s)}}$ are origin district fixed effects, and X are the same set of controls as in Equation 10.

The instrument induces 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, green, and orange zones already traded intensively with sellers in certain lockdown zones, and there is heterogeneity in responses. For instance, if buyers in red zones traded mostly with sellers in red zones, then our instrument may estimate effects on firms induced by having more red-zone sellers, and so upweight effects on buyers in red zones. Figure 2 shows that this is unlikely to be the case: In general firms from red, orange, and green zones had similar interactions with firms from red, orange, and green zones.

We also consider whether certain industries source intensively from firms located in certain zones. For instance, if all the rubber supply of firms in this production network comes from red-zone suppliers, then buyers of rubbers will have a hard time finding substitutes. Once again, if there is heterogeneity in responses by industry, our estimate LATE elasticity would weight the rubber industry higher than non-rubber industries. While not a source of bias, it does affect the interpretation of the estimated parameter. In figures 6 and 7, we plot the shares of total purchases of each industry that are sourced from firms in red, orange, and green zones. With the exception of HS industry 19 (arms and ammunitions) we see that there is no noticeable degree of concentration of suppliers from any particular color zone.

6.3 Estimation results

In this section we show results of the estimation. First, we report OLS estimates in Table 2. The implied elasticities exhibit a robust value of 0.77 across all the different specifications. This suggests that at the firm level, suppliers act as complements rather than substitutes for a buyer. From Equation (8) we can see that, once we take into account second order effects, an elasticity of substitution less than 1 implies that the aggregate impact of negative shocks are amplified, while the positive ones get dampened.

Nevertheless, as we describe in the previous section, it is very likely that our estimators are contaminated by simultaneous demand shocks that happened during Covid-19. We now report our results using our proposed instruments. Our 2SLS estimators are reported in Table 3. Across specifications, our estimators exhibit a robust value of around 0.38. As discussed in section 6.2, the bias is in the expected direction if we expect the Covid-19 shock to also induce negative demand shocks, thereby over-estimating ϵ. Also, the Kleibergen-Paap rk Wald F statistic shows that our instruments are strong across all specifications. All this implies an even higher degree of complementarity between suppliers than implied by our OLS estimates. This further reinforces the role for this elasticity to amplify negative shocks.

Finally, we also show our estimators at the industry level. We consider the classification of *Section* in India.[10] We show our results in Table 6. We do not report the elasticities for industries where we lack significant observations across zones, or have weak first-stages in the case of the IV regression. In terms of the OLS estimators, we can see that estimators are around the overall value of $\epsilon = 0.77$, but they are all between 0 and 1. This suggests that, even though all these industries exhibit complementarity between suppliers, the extent of this

[10] https://gstportalindia.in/list-of-hsn-code-chapter-wise-updated/

complementarity varies. Finally, our 2SLS estimators using the seller-level instrument shows an interesting picture. In most cases, the implied elasticities are corrected downwards, reflecting possibly negative demand shocks, and thus an even higher degree of complementarity for most industries. Interestingly, there are some industries that now exhibit substitutability such as processed foods, textiles, and the metal industry. This estimation results show that while in the overwhelming majority of industries suppliers act as complements in the production process, there are some industries where inputs from different suppliers could act as substitutes, and we should be mindful of heterogeneity across industries in understanding how shocks propagate through supply chains. In industries where suppliers are substitutes, a negative productivity shock to a single supplier would lead to a less severe shock propagation through supply chains compared to industries where the suppliers are complements. In future versions of the paper, armed with estimates of these firm level elasticities of substitution, we will use equation 8 to study how shocks propagate through the economy when there are supply shocks to different firms and industries.

6.4 The effect of aggregation on firm-level elasticities of substitution

In section 6.3 we establish that suppliers of the same HS-2 product are highly complementary in production. In this section we analyze whether the choice of aggregation has any effect on this estimate. First, we re-estimate the elasticities using the HS-4 level of aggregation. This further restricts the set of suppliers that a firm can substitute from. Table 4 reports the results for our IV estimation. We find that this estimate of the HS-4 digit elasticity is slightly lower at 0.286, compared to the estimated elasticity at the 2 digit level. However, the estimated elasticity at the 4-digit level lies within the confidence interval of the estimated elasticity at the 2-digit level. In table 5 we further show the IV results using the HSN classification. This exercise expands the set of suppliers a firm can source from, compared to the above two cases when we restrict the firms to source from suppliers within the same HS-2 or HS-4 product. The estimated value of this elasticity rises to 0.581. However, looking across the specifications, we find that all the values still lie within the confidence interval of the estimate from our main HS-2 specification. From this exercise, we can conclude that at the very short-run, inputs are highly complementary, whether we measure this elasticity across suppliers of the same HS section, HS-2, or HS-4 product.

6.5 Preliminary simulations

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. As a first step, we use production network data from March 2019 to February 2020, and randomly sample less than 1% of this data to carry out preliminary simulations given computational constraints. In future drafts we plan to use the whole production network data for this exercise. In this randomly sampled limited production network data, there are 6569 firms, and the average firm only supplies to about 2 other firms. The firm with the highest number of connections supplies to about 10 other firms. Today, the point of this exercise is only to show the importance of these elasticities.

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 from 2-digit HSN industries, (ii) firm sales to 2-digit HSN industries, (iii) labor employed by each firm, and (iv) final sales by each firm. The first two submatrices are directly constructed from the data. For today's exercise, we have assumed that the labor employed and the final goods sold by each firm is the same, and is equal to 1.

Given the values of the elasticity of substitution between different consumption goods set at $\sigma = 4$, the substitution across different intermediate bundles set at $\zeta = 0.2$, we vary the elasticity of substitution ϵ across different suppliers of the same HS-2 good from .01 to 2 to see how that affects GDP. In this preliminary simulation, we shock the productivities of 3 randomly chosen firms by -1%. In figure 8 we then plot the relative fall in GDP in this imaginary economy consisting of the 6569 randomly selected firms where $\epsilon = (0.001, 0.17, 0.38, 0.59, 1, 1.21, 1.42)$. $\epsilon = 0.38$ is our main estimate, and we consider intervals around 0.38 depending on the confidence interval of our estimates (0.21). $\epsilon = .001$ represents a small enough elasticity (close to zero) and $\epsilon = 1$ represents the case where expenditure shares do not vary at all with input prices, ie, the Cobb Douglas case. We find that even when we perturb the productivity of a small number of firms (.04% of firms in the economy) by 1%, there is a larger effect of the shock on GDP when the inputs are complementary compared to the Cobb-Douglas case (about -0.0002%). As suppliers become more and more substitutes, the negative second order effect of the shock on GDP reduces. Our simulation results echo the findings of Baqaee and Farhi (2019) who show that a 13% shock to the productivity of the oil industry reduced aggregate output by about 0.61% more compared to the case when there are no second order effects, amplified by input complementarity. This toy example with a small number of firms from our data conveys our main message: It is important to know the extent of complementarity or substitutability between suppliers in order to understand how shocks propagate through firm networks. We provide novel estimates of these elasticities across suppliers of the same product at the firm level and show that these estimates are robust to the choice of product aggregation. In future versions of this paper, we will first repeat this exercise with the full data on production network. We will then conduct counterfactuals where we will examine given the production network data and the knowledge of firm level elasticities of substitution, what would be the effect on GDP if government policies would allow goods to pass even in red zones depending on how connected a single supplier is and given the importance of its input in the production network.

7 Conclusion

In this paper, we leverage variation in input prices derived from quasi-exogenous variation following the Covid-19 lockdown in India, and provide one of the first estimates of elasticities of substitution across suppliers within the same industry. We find that, even at this very granular level, inputs are highly complementary. Our findings have important implications for the propagation of shocks through firm production networks. Since inputs are complementary, even negative shocks to small subsets of firms that are highly linked in the supply chains can have large negative effects on the aggregate economy by propagating through firm networks. Our paper also provides evidence that domestic lockdown policies undertaken during the Covid-19 crisis in India, also fairly common in many countries all over the world during the crisis, have severely reduced the number of transactions and increased the costs of inputs as measured by unit values. This evidence, combined with the fact that inputs, even at the supplier level are highly complementary, could explain at-least part of the huge GDP decline in India observed during the Covid-19 crisis of 2020. In subsequent iterations, we will show how important these firm-level elasticities are in explaining this decline in GDP, and how much GDP decline we would be able to explain if we do not take into account these elasticities.

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.

Atalay, E. (2017). How important are sectoral shocks? *American Economic Journal: Macroeconomics 9*(4), 254–80.

Baldwin, R. and E. Tomiura (2020). Thinking ahead about the trade impact of Covid-19. *Economics in the Time of Covid-19 59*.

Baqaee, D. and E. Farhi (2020). Nonlinear production networks with an application to the covid-19 crisis. Technical report, National Bureau of Economic Research.

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.

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 (May 29, 2021)*.

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.

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), 134–146.

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. Technical report, National Bureau of Economic Research.

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), 1255–1321.

Demir, B., A. C. Fieler, D. Xu, and K. K. Yang (2021). O-ring production networks. Technical report, National Bureau of Economic Research.

Demir, B. and B. Javorcik (2020). Trade finance matters: evidence from the Covid-19 crisis. *Oxford Review of Economic Policy*.

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? Technical report, National Bureau of Economic Research.

Fajgelbaum, P. D. and E. Schaal (2020). Optimal transport networks in spatial equilibrium. *Econometrica 88*(4), 1411–1452.

Feenstra, R. C., P. Luck, M. Obstfeld, and K. N. Russ (2018). In search of the Armington elasticity. *Review of Economics and Statistics 100*(1), 135–150.

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. et al. (2020). How did china's covid-19 shutdown affect us supply chains? Technical report, Federal Reserve Bank of New York.

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), 1597–1629.

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.

Lafrogne-Roussier, R., J. Martin, and I. Méjean (2021). Supply shocks in supply chains: Evidence from the early lockdown in china. Technical report, Mimeo, CREST.

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

Panigrahi, P. P. (2021). *Endogenous Spatial Production Networks: Theory, Estimation, and Evidence from Indian Firm-to-Firm Linkages*. Ph. D. thesis, University of California, Berkeley.

Peter, A., C. Ruane, et al. (2020). The aggregate importance of intermediate input substitutability. *Princeton University, memo*.

Ravindran, S. and M. Shah (2020). Unintended consequences of lockdowns: Covid-19 and the shadow pandemic. Technical report, National Bureau of Economic Research.

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), 503–560.

Tashereau-Dumouchel, M. (2020). Cascades and fluctuations in an economy with an endogenous production network. *Available at SSRN 3115854*.

TABLES

TABLE 1: Summary statistics

	Panel A: 2019		
	Jan-March	April-June	July-September
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 is comprised by two panels. Panel A contains information about number of sellers, buyers, transactions, and total sales for periods January-March, April-June, July-September for year 2019. Panel B is the same as Panel A, but with respect to 2020.

TABLE 2: OLS, 2-digit HSN

	(1)	(2)	(3)	(4)	(5)
$\log\left(\frac{\hat{p}}{\hat{\hat{p}}}\right)$	0.23	0.23	0.23	0.23	0.23
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.271	0.272	0.272	0.271	0.272
Obs	4450688	4450656	4449474	4450663	4449416
ϵ	0.769	0.769	0.769	0.769	0.769
HSN × Month FE		Y			Y
Destination × Month FE			Y		Y
Origin				Y	Y

Notes: These are the estimators resulting from running OLS regressions from equation (10). An industry is defined as 2-digit HSN codes. The first row reports the estimator associated to log relative unit values. Standard errors are clustered at the origin/destination state level, and are reported in parentheses below each estimator. The fifth row reports the implied value for ϵ, which is 1 minus the estimator on the first row. The table contains 5 columns, one for each set of fixed effects. The last three rows denote which fixed effects are included. All specifications include the controls mentioned in the paper.

TABLE 3: 2SLS, 2-digit HSN

	OLS	(2)	(3)	(4)	(5)	(6)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.230	0.622	0.622	0.616	0.586	0.622
	(0.006)	(0.214)	(0.234)	(0.132)	(0.306)	(0.217)
Obs	4449449	4449449	4449449	3213758	3213758	4449449
K-PF		17.026	16.958	114.7503	74.811	16.958
J-stat		3.082	2.906	2.929	3.070	2.906
ϵ	0.770	0.377	0.377	0.383	0.413	0.377
Instrument I		Y	Y	Y	Y	Y
Instrument II				Y	Y	

Notes: These are the estimators resulting from running 2SLS regressions. Industries are 2-digit HSN codes. The second stage corresponds to equation (10). The first stage using the seller-level instrument corresponds to equation (11), while the one using seller/buyer corresponds to equation (12). Each column corresponds to the different combinations of instruments which we report on the last two rows. The difference between columns (2) and (6) is that, in the latter, standard errors are calculated under bootstraping with 50 iterations. The first row reports the estimator associated to log relative unit values. Standard errors in columns (1), (2), (4), and (5) are clustered at the origin/destination state level, while columns (3) and (6) are clustered at the origin level. Standard errors are reported in parentheses below each estimator. The fourth row reports the Kleibergen-Paap rk Wald F statistic for instrument weakness. The fifth row reports the J statistic for overidentification. The sixth row reports the implied value for ϵ, which is 1 minus the estimator on the first row. All specifications include the controls mentioned in the paper.

TABLE 4: 2SLS, 4-digit HSN

	(1)	(2)	(3)	(4)	(5)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.713	0.713	0.531	0.600	0.713
	(0.313)	(0.329)	(0.158)	(0.438)	(0.362)
Obs	5478629	5478629	3945976	3945976	5478629
K-PF	8.817	8.608	81.811	12.634	8.608
J-stat	0.054	0.065	0.549	2.930	0.065
ϵ	0.286	0.286	0.468	0.399	0.286
Instrument I	Y	Y	Y	Y	Y
Instrument II			Y	Y	

Notes: These are the estimators resulting from running 2SLS regressions. Industries are 4-digit HSN codes. The second stage corresponds to equation (10). The first stage using the seller-level instrument corresponds to equation (11), while the one using seller/buyer corresponds to equation (12). Each column corresponds to the different combinations of instruments which we report on the last two rows. The first row reports the estimator associated to log relative unit values. Standard errors in columns (1), (3), and (4) are clustered at the origin/destination state level, while columns (2) and (5) are clustered at the origin level. The difference between columns (2) and (5) is that, in the latter, standard errors are calculated under bootstraping with 50 iterations. Standard errors are reported in parentheses below each estimator. The fourth row reports the Kleibergen-Paap rk Wald F statistic for instrument weakness. The fifth row reports the J statistic for overidentification. The sixth row reports the implied value for ϵ, which is 1 minus the estimator on the first row. All specifications include the controls mentioned in the paper.

TABLE 5: 2SLS, HSN sections

	(1)	(2)	(3)	(4)	(5)
$\log\left(\frac{\hat{p}}{\tilde{\hat{p}}}\right)$	0.418	0.418	0.507	0.644	0.418
	(0.281)	(0.305)	(0.128)	(0.362)	(0.206)
Obs	3870856	3870856	2799889	2799889	3870856
K-PF	28.169	31.042	25.610	17.814	31.042
J-stat	2.379	1.868	2.562	5.536	1.868
ϵ	0.581	0.581	0.492	0.355	0.581
Instrument I	Y	Y	Y	Y	Y
Instrument II			Y	Y	

Notes: These are the estimators resulting from running 2SLS regressions. Industries are HSN sections. The second stage corresponds to equation (10). The first stage using the seller-level instrument corresponds to equation (11), while the one using seller/buyer corresponds to equation (12). Each column corresponds to the different combinations of instruments which we report on the last two rows. The first row reports the estimator associated to log relative unit values. Standard errors in columns (1), (3)and (4) are clustered at the origin/destination state level, while columns (2) and (5) are clustered at the origin level. The difference between columns (2) and (5) is that, in the latter, standard errors are calculated under bootstraping with 50 iterations.Standard errors are reported in parentheses below each estimator. The fourth row reports the Kleibergen-Paap rk Wald F statistic for instrument weakness. The fifth row reports the J statistic for overidentification. The sixth row reports the implied value for ϵ, which is 1 minus the estimator on the first row. All specifications include the controls mentioned in the paper.

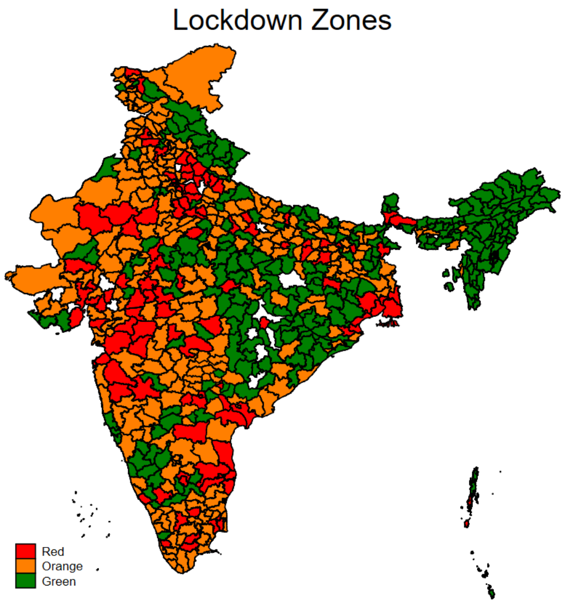
TABLE 6: Elasticities by industry

Section	Name	OLS elast.	2SLS elast.
2	Vegetables	.8082675	.
3	Fats and Oils	.825045	.3875365
4	Processed foods	.7458998	1.141
5	Minerals	.8220726	.5755809
7	Plastics	.8097205	.
9	Wood	.8905362	.
10	Wood derivatives	.8832779	.8700905
11	Textiles	.8635682	1.636
12	Clothing	.8435352	.3459941
13	Handcrafts	.778517	.
15	Metal	.8466598	1.165
16	Machinery	.6709916	.
17	Transport equipment	.5481665	.216569
18	Surgical instruments	.6465395	.
21	Art	.7354167	.7348618

Notes: Each row corresponds to a Section, which is a conglomerate of HSN codes. The second column mentions the name of the section. The third column reports the implied elasticity by OLS estimation as in Equation (10). The fourth column reports the implied elasticity by 2SLS estimation as in Equation (10) for the second stage, and using seller-level instruments for the first stage as in Equation (11). Missing elasticities were not able to be estimated due to lack of statistical power, or the F-stat reflected weak instruments so they were not reported. All regressions for each section include all controls as described in the paper, destination state/month and origin state fixed effects, and standard errors are clustered at the origin/destination state level.

Figures

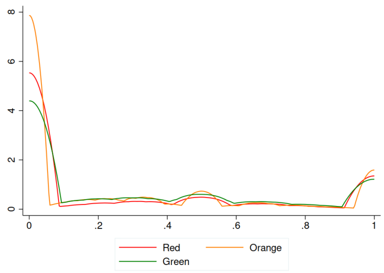
Figure 1: Map showing India's lockdown zones



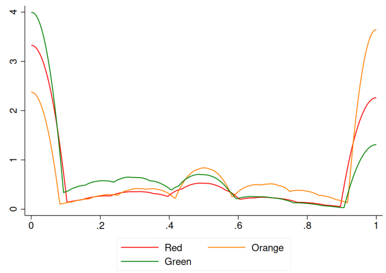
Notes: The map above shows the lockdown zones across Indian districts, where the lockdown was announced on March 25,2020.

FIGURE 2: SHARE DISTRIBUTIONS OF COLORS

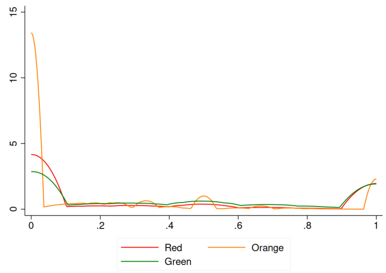
(a) Sellers in Red



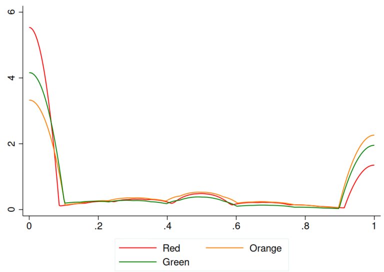
(b) Sellers in Orange



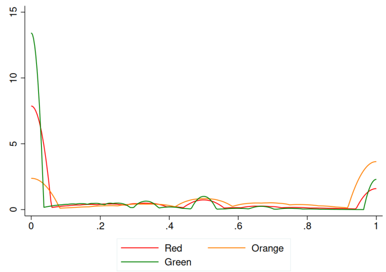
(c) Sellers in Green



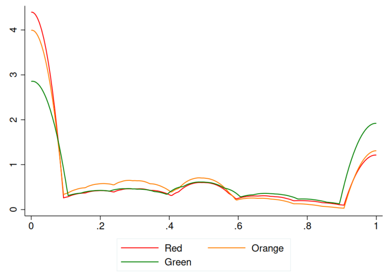
(d) Buyers in Red



(e) Buyers in Orange



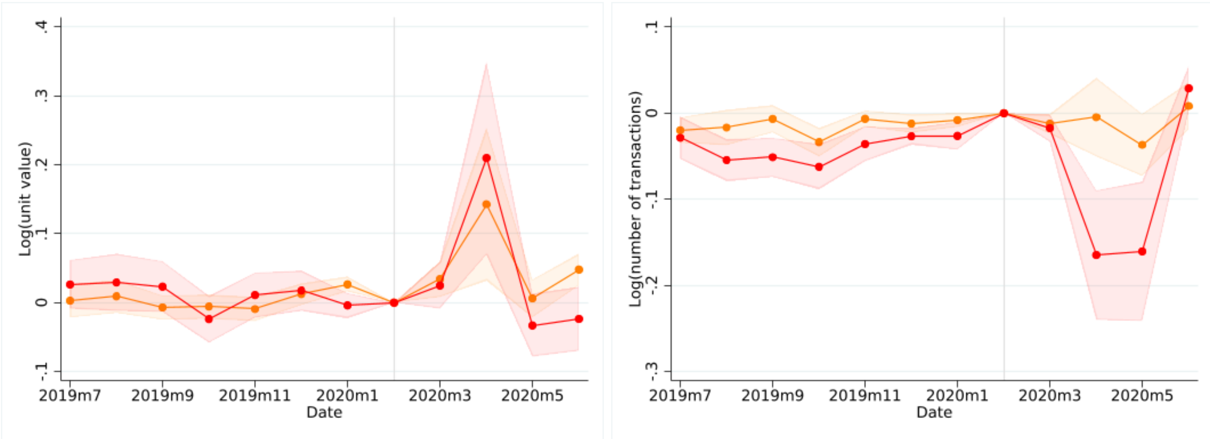
(f) Buyers in Green



Notes: In the three upper panels, each panel plots the distribution of the share of buyers located in *Red*, *Orange*, or *Green* districts. Each panel corresponds to sellers located in their corresponding color district. In the lower three panels, each panel plots the distribution of the share of sellers located in *Red*, *Orange*, or *Green* districts. Each panel corresponds to buyers located in their corresponding color district. The time period is April 2018 - February 2020.

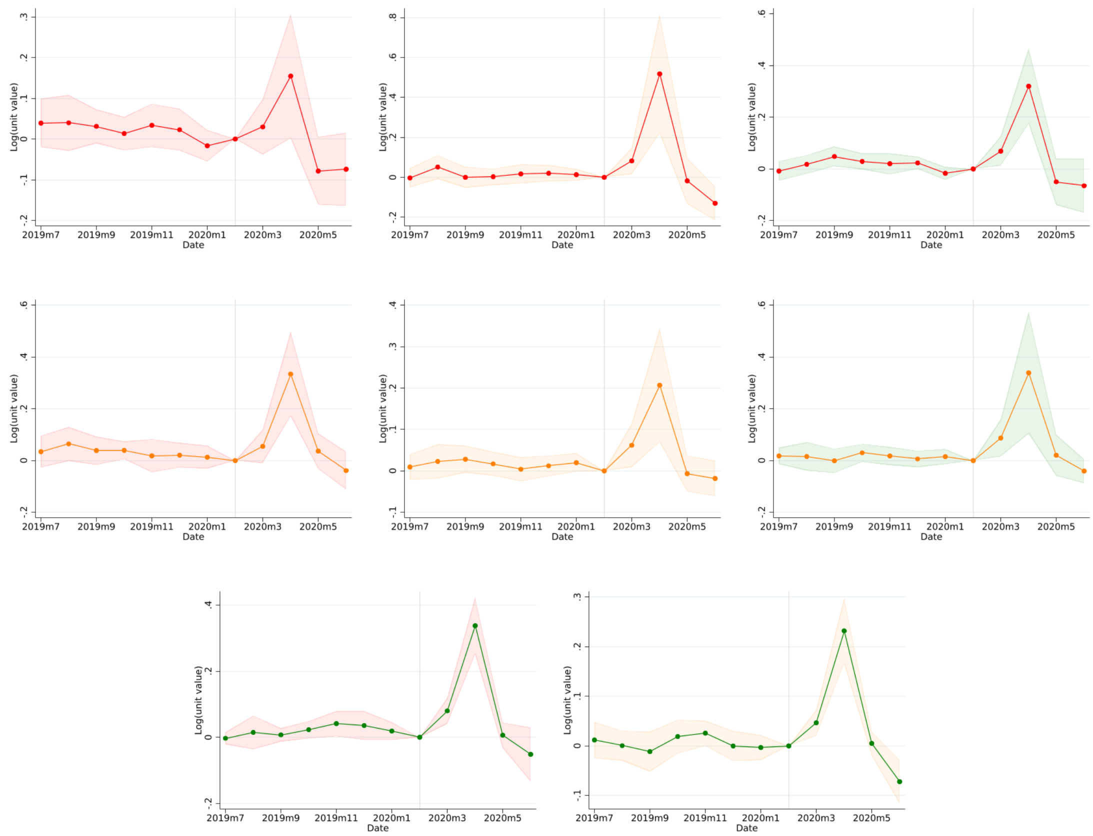
FIGURE 3: SELLER REGRESSIONS

(a) Unit value (b) Number of transactions



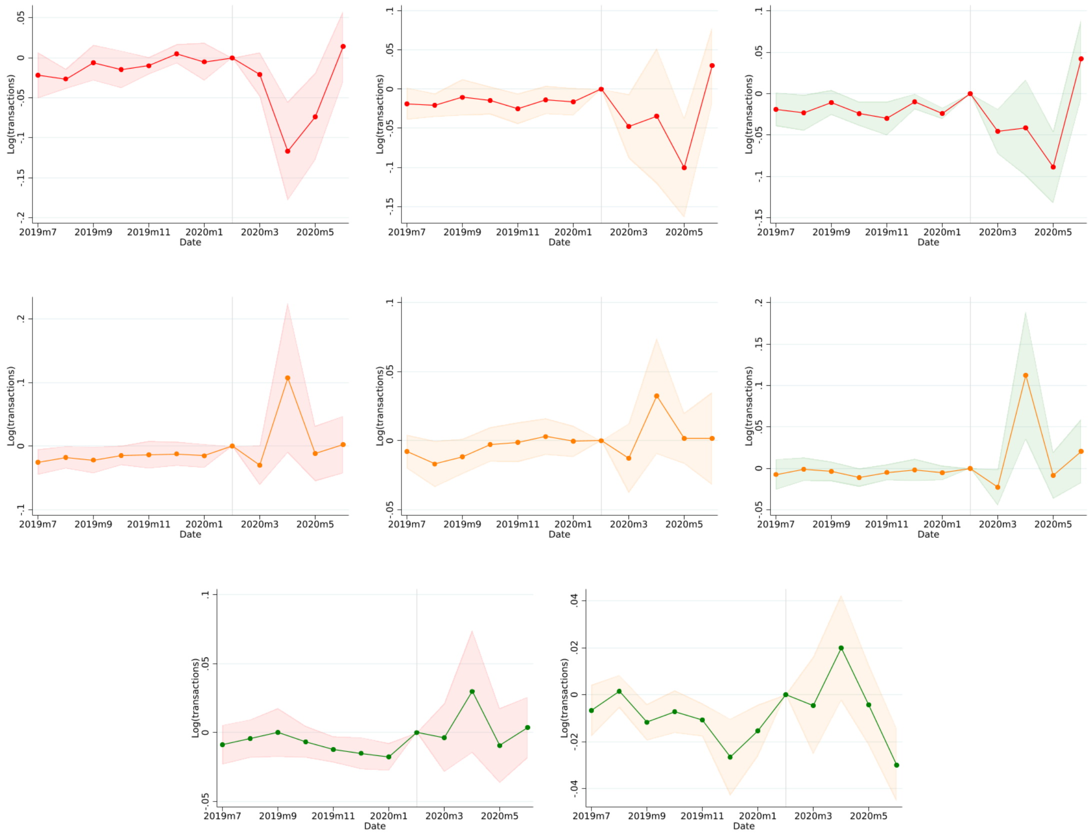
Notes: In each plot, the horizontal axis is month, and the vertical one is the estimator of interest as in Equation (1) for each month. The title on each plot denotes the outcome of interest in logs. Regressions include district, month , and HSN fixed effects. Standard errors are clustered at the district level. All controls mentioned in the paper are included. The vertical line in February 2020 splits pre and post lockdown periods. The baseline category are sellers located in *Green* districts in February 2020. The *Red* and *Orange* lines plot the percentage changes in unit values and number of transactions for sellers located in *Red* and *Orange* districts, respectively. The shaded area are confidence intervals.

FIGURE 4: UNIT VALUE, SELLER-BUYER REGRESSIONS



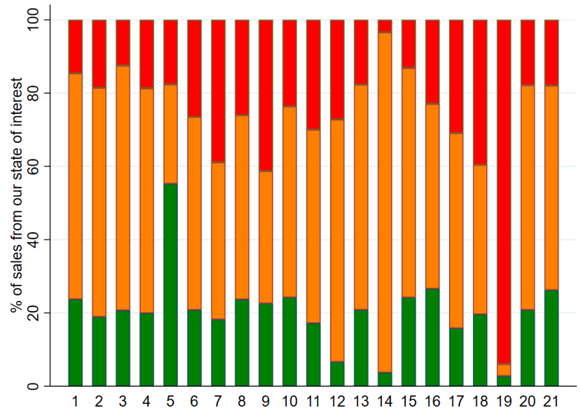
Notes: In each plot, the horizontal axis is the month, and the vertical one is the estimator of interest associated to log unit values as in Equation (2) for each month. Regressions include HSN fixed effects, and standard errors are clustered at the origin and destination district level. All controls mentioned in the paper are included. The vertical line in February 2020 splits pre and post periods. The baseline category are sellers and buyers located in *Green* districts on February 2020. The color of the line denotes the color of the district the seller is located, while the color of the shaded confidence interval denotes the color o the district the buyer is located.

FIGURE 5: NUMBER OF TRANSACTIONS, SELLER-BUYER REGRESSIONS



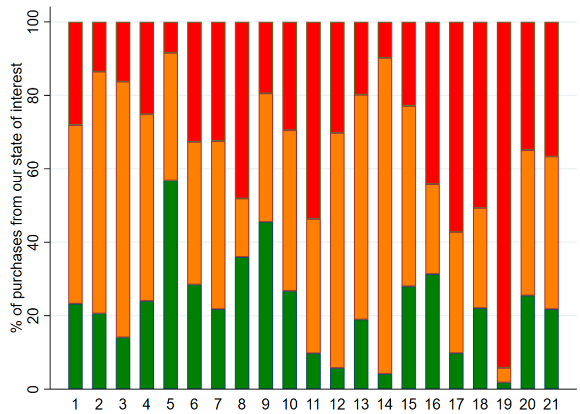
Notes: In each plot, the horizontal axis is the month, and the vertical one is the estimator of interest associated to log number of transactions as in Equation (2) for each month. Regressions include HSN fixed effects, and standard errors are clustered at the origin and destination district level. All controls mentioned in the paper are included. The vertical line in February 2020 splits pre and post periods. The baseline category are sellers and buyers located in *Green* districts on February 2020. The color of the line denotes the color of the district the seller is located, while the color of the shaded confidence interval denotes the color o the district the buyer is located.

FIGURE 6: % OF SALES, BY COLOR OF DESTINATION DISTRICTS



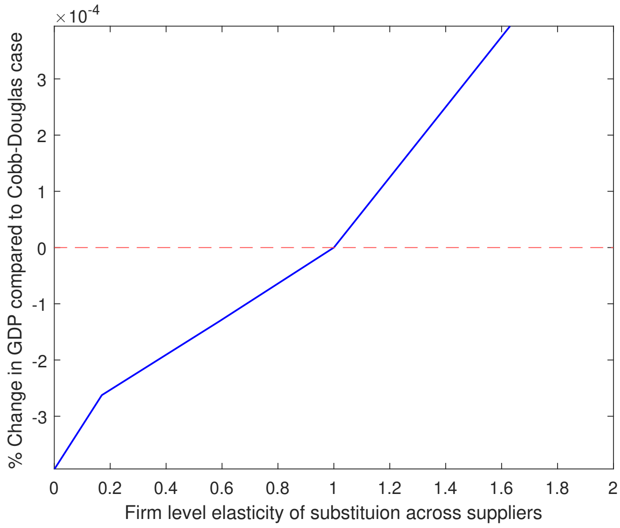
Notes: For each HSN section (horizontal axis), we plot the share of total purchases of firms located in our state by color of selling districts.

FIGURE 7: % OF PURCHASES, BY COLOR OF ORIGIN DISTRICTS



Notes: For each HSN section (horizontal axis), we plot the share of total sales of firms located in our state by color of buying districts.

FIGURE 8: % CHANGE IN GDP FROM NEGATIVE PRODUCTIVITY SHOCK FOR DIFFERENT VALUES OF ELASTICITY ACROSS SUPPLIERS



Notes: For each firm-level elasticity of substitution across suppliers of the same HS-2 product (horizontal axis), we plot the % change in GDP compared to the Cobb-Douglas case (elasticity=1) after considering the second order effects of a negative productivity shock of 1% to 3 randomly selected firms out of all firms.

Appendix for online publication only

A Data

Exposure variables. We have two exposure variables: $ED_{si,t}$ and $IM_{si,t}$. The first one denotes the exposure of firm s selling product i to global demand shocks in month t. The second one denotes 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 the beginning of the sample, $Y_{si,m,0}$ is the value of goods of seller s of product i shipped from country m in the beginning of the sample, $X_{i,x,t}$ 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 do a weighted sum of these measures across countries, such that

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

B Dijkstra algorithm

Here we describe the implementation of the Dijkstra algorithm we use to construct the seller/buyer instrument.

1. We obtain shapefiles of district administrative boundaries for India according to India's 2011 census;

2. We reprojected shapefiles into the Asian/South Equidistance Conic projection so we minimize the distortion into the distances between districts;

3. We obtain the location of the centroid of each district in India;

4. For each centroid, we obtain the k closest centroids according to Euclidean distances. Notice that we can't consider the whole distance matrix to find least-distance routes since, in that case, the least-distance route between any pair of centroids is just the Euclidean distance between them. Then, we follow Fajgelbaum and Schaal (2020) and consider

$k = 8$ such that we consider all possible directions. We then have a well-defined network structure we can use to feed into the Dijkstra algorithm;

5. We run the Dijkstra algorithm. For all district pairs, the algorithm provides us with the list of all districts that comprise the route between district pairs, and the distance for each leg that comprise the route;

6. Using the name of the districts, we use the lockdown data to assign a lockdown color to each district along the route. We can the calculate the share of districts in a route belonging to red, yellow, or green colors. When calculating these shares, we rule-out the zone where the buyer resides to clean from demand-side shocks. Using also the distance of each leg, we also calculate the share of all the route that belongs to red, yellow, or green. A "red" leg is one where the origin district was red. Here we also ignore the color of the district where the buyer resides.

C Derivations

C.1 Expression to estimate firm-level elasticities of substitution

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

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

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

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

where $(\dots)$ are components that are not of interest to us since they will cancel anyway. 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 $PM_{i,bj} \equiv \sum_{s'} PM_{s'i,bj}$. Finally, recall this is the basic equation we take to the data. In the next section we modify this expression to address endogeneity issues.

C.2 Addressing unobservable productivity shocks

In this section we derive the expressions that allows us to construct price indexes based on observable data, which we then use to obtain estimators of our elasticities of substitution. First, throughout the derivation in Appendix C.1, consider the following step:

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

In the data we observe the production network over time, so we can introduce a time dimension such that

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

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 $\widehat{x}_t \equiv \frac{x_t}{x_{t-1}}$ is the change of a variable x_t. 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 taste shocks of buyers for different sectors are time invariant. Intuitively, the taste of a shoe-maker for different suppliers of leather can change over time, but not its overall preference for leather to make shoes. We can rewrite this expression as

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

where $s_{si,bj,t} = \frac{PM_{si,bj,t}}{PM_{i,bj,t}}$. Notice that this equation holds for all s, so we can apply a geometric mean to this expression, such that

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

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

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

Then, the last term of our expression is

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

So our final expression boils down to

$$\hat{p}_{i,bj,t}^{1-\epsilon} = \frac{\hat{\tilde{p}}_{si,bj,t}^{1-\epsilon}}{\hat{\tilde{s}}_{si,bj,t}}$$

where $\tilde{p}_{i,bj,t} \equiv \prod_s p_{si,bj,t}^{\frac{1}{N_{i,bj,t}}}$ is a geometric mean across suppliers of unit values, and $\tilde{s}_{i,bj,t} \equiv \prod_s s_{si,bj,t}^{\frac{1}{N_{i,bj,t}}}$ is a geometric mean across suppliers of expenditure shares, $s_{si,bj,t} \equiv \frac{PM_{si,bj,t}}{PM_{i,bj,t}}$, and $N_{i,bj,t}$ is the number of suppliers that firm sourced from in time t.[11] Notice that we have reached to our objective, since now $\widehat{\tilde{p}}_{i,bj,t}$ is independent of productivity shock and it depends on variables directly observed in the data. We now have

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

C.3 *Addressing endogeneity concerns*

The equation from previous section is what we take to the data. Nevertheless, there are further endogeneity issues that would contaminate our estimators 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 estimators 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

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

In the data what we actually observe are transactions between sellers and buyers, where the good being traded belongs to a certain HSN sector. Then, we assume that the sector of the good solely determines the sector the seller. Now, we can then express this in changes, such that

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

In logarithms, we have

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

Now we explain how we construct the instrument at the seller/buyer level. We have to take a stance in the functional form of the trade cost τ. We assume that trade costs are an

[11]This is a simplified version of the main result by Redding and Weinstein (2020) where we are not considering entry/exit notions.

proportional to the travel time of the transportation of intermediate inputs, such that

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

If we express this in changes, we get

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

We exploit variation from the Covid 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:

$$\widehat{TravelTime}_{sb,t} = \exp\left(\gamma^R Red_{o_{(s)}} Lock_t + \gamma^O Orange_{o_{(s)}} Lock_t + \nu_{si,bj,t}\right),$$

where $Red_{o_{(s)}d_{(b)}}$ and $Orange_{o_{(s)}d_{(b)}}$ are the share of 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 A. X are included instruments. $Lock_t$ is a dummy variable that equals 1 for April and May 2020 (i.e. the treatment period), which is when the lockdown was implemented, 0 otherwise. $\nu_{si,bj,t}$ is an unobservable error term. If we combine all these expressions, we can get our first stage regression as

$$\begin{aligned}
\log\left(\widehat{p}_{si,bj,t}\right) &= \log\left(\widehat{\tau}_{sb,t}\right) + \log\left(\widehat{p}_{i,t}\right), \\
&= \log\left(\widehat{TravelTime}_{sb,t}^{\sigma}\right) + \log\left(\widehat{p}_{i,t}\right), \\
&= \sigma \log\left(\widehat{TravelTime}_{sb,t}\right) + \log\left(\widehat{p}_{i,t}\right), \\
&= \sigma \log\left(\exp\left(\gamma^R Red_{o_{(s)}} Lock_t + \gamma^O Orange_{o_{(s)}} Lock_t + \nu_{si,bj,t}\right)\right) + \log\left(\widehat{p}_{i,t}\right), \\
&= \sigma\left(\gamma^R Red_{o_{(s)}} Lock_t + \gamma^O Orange_{o_{(s)}} Lock_t + \nu_{si,bj,t}\right) + \log\left(\widehat{p}_{i,t}\right), \\
&= \sigma\gamma^R Red_{o_{(s)}} Lock_t + \sigma\gamma^O Orange_{o_{(s)}} Lock_t + \log\left(\widehat{p}_{i,t}\right) + \sigma\nu_{si,bj,t}.
\end{aligned}$$