Production Networks and Firm-level Elasticities of Substitution*

Brian Cevallos Fujiy† Devaki Ghose‡ Gaurav Khanna §

This version: May 2022

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-4 industry. This estimate is particularly relevant for the transmission of supply shocks. If suppliers in the lockdown and non-lockdown zones are complements rather than substitutes in production, this shock can amplify by further transmitting downstream and upstream through the supply chain. We find that even at this very granular supplier level, inputs are highly complementary, with an estimated elasticity 0.58. 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. We then use our elasticities and simulate the impact of the Covid-19 lockdowns to find that under our estimated elasticities the overall fall in output is substantial and widespread.

Keywords: shock propagation, resilience, Covid-19, networks
JEL Codes: F41, F44

*We thank Ahmed Magdy and 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 Mid-Atlantic Trade Conference 2022, MIEDC 2022, joint Banque de France, University of Surrey, CEPR, and World Bank workshop on globalization 2022, IMF, the Bank of Italy, STEG 2022, Lund University, World Bank, and Ana Cecilia Fieler, Pablo Garriga, Carolina Sanchez, and Ludovic Panon for insightful discussions. 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 elasticities 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-4 industry but across different suppliers are highly complementary, with an estimated value of 0.58. 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 0.36 to 0.58. These estimates are slightly higher than Atalay (2017) who estimates this to be around 0.1, 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-4 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.[3]

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, 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.[4]

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 industries. 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

[3] We also estimate this elasticity at a more aggregate level, that is, the firm-level elasticity of substitution across different industries, and find evidence of complementarity across industries, in line with the findings of Atalay (2017).

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

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 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-4 product from different suppliers to the equilibrium relative prices. That is, within each HS-4 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 0.58. 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.

Third, 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. We find that a 15% productivity shock to firms in the red zone reduces GDP by 6.05%. This fall would be 0.45 pp less in a model where firms in the same HS-4 industry are considered substitutes ($\epsilon = 1.75$), and 0.22 pp more when firms in the same HS-4 industry are considered almost leontieff ($\epsilon = .001$) . In terms of GDP losses, given that the quarterly GDP of this state was close to 32.5 billion USD in 2020-2021, the additional losses due to firm-level complementarities translate into 146.2 million USD (4.2 USD per capita), compared to the case when firms are substitutes. Next we investigate whether aggregate GDP losses in the face of large productivity shocks are less if policy makers let large firms (as measured by the final sales) or more connected firms (as measured by the leontief inverse which measures the direct and indirect linkages) to operate. We show that as the level of complementarity and the magnitude of the negative productivity shock increase, it pays more to save the more connected firms. A lot of importance, both in policy and in academic circles, has been paid to large firms since Hulten's 1978 theorem which emphasized the importance of firm sizes in the propagation of shocks through production networks. We show that in the face of large negative productivity shocks and high levels of complementarity across suppliers, the more connected firms could become more important than large firms in the propagation of shocks through firm networks.

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 the values associated with 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-4 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."[5] 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; Bas et al., 2022; Chakrabati 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 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

In this section we provide a general overview of our data and context.

Firm-to-firm trade. Our primary data source is daily establishment level transactions.[6] 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

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

[6] 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.

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 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 4-digit HSN/month/transaction level, and then construct implied unit values. We can then collapse the data at the 4-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 we map 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.[7]

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).

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`.[8] For each firm, we construct two variables that measure the firm's exposure to global demand and supply shocks that vary at the HS-4 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.

Before using the variations in the lockdown to understand how firm to firm transactions are affected, we verify the stringency of these lockdowns in figure 4 using google mobility data. The data shows how the number of visitors to (or the time spent in) categorized places change compared to baseline days. The baseline day is the median value from the 5-week period Jan

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

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

3 – Feb 6, 2020.[9] As is clear from the graph, until March 2020, there were essentially no differences in mobility trends across red, orange, or green zones. But starting in April 2020, we see that there is a substantial reduction in different types of activities (spending time in retail and recreation, grocery and pharmacy, parks, commuting, and workplaces) in red zones compared to green zones. While such differences also exist between orange and green zones, these differences are most stark between red and green districts. People in red zones also spend more time at home compared to people in either orange or green zones. We notice that starting August 2020, a few months after the centralized lockdown was over, these differences start to reduce, and by December 2020 these differences, especially in workplace mobility, becomes very small. We therefore only restrict our sample to June 2020, close to the period the centralized lockdown was effective.

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 March-May 2020 to estimate the effect of lockdown policies on prices and quantities of intermediate inputs between firms.[10]

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 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.

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

[10] 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).

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-4 industry i in month t, ι_i are 4-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. Standard errors are clustered at the seller origin district level.

4.1.2. Seller-buyer level regressions

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

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

where $Y_{si,b,t}$ are unit values or number of transactions in logs between seller s in HS-4 industry i and a buyer b in month t. $\delta_{o_{(s)}}$, $\delta_{d_{(b)}}$, δ_i, and δ_t are origin, destination, industry, 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 $\gamma_{d_{(b)}}^z$ 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_t^{xz}. Our base time category is February 2020 which is just before lockdowns began. Standard errors are clustered at the seller-buyer origin-destination district level.

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 5a 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 6, 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 6 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 6, 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 20 pp to 30 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-HS4 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 5a, 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 5b 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 7, 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 7a, 7b, 7c, we see that the fall in the transactions only increases in the severity of the lockdown at the seller's district. From figure 5b, 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 use the model to simulate the effects of negative productivity shocks on GDP. We follow the framework in Baqaee and Farhi (2019) to write out a nested CES economy in standard form.[11]

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

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

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

[11] 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.

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:

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

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 industry 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-4 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 $\widehat{\Omega}$ 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 $\widehat{\theta}$, where $\widehat{\theta} = (\alpha, \zeta, \epsilon, \sigma)$.

Formally, a nested-CES economy in standard form is defined by $(\widehat{\Omega}, \widehat{\theta})$. What distinguishes factors from goods is that factors cannot be produced. The $(2+N+I) \times (2+N+I)$ input–output matrix $\widehat{\Omega}$ 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 $d\log A_j < 0$.

The second order term captures the reallocation effect: In response to a negative shock to industry j, all industries k that are downstream of j may readjust their demand for all other inputs. Crucially, the impact of such readjustments by any given k on the output of industry i depends on the size of industry k as captured by its Domar weight λ_k, the elasticity of substitution σ_k in $k's$ production function, and the extent to which the supply chains that connect i and j to k coincide with one another, as given by the covariance term.

5.1 Deriving the estimating equation for the firm-level elasticity of substitution across firms

Using the model outlined above, in this section we derive the firm-level elasticity of substitution across suppliers. 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 industries J. A seller is denoted by $s \in F$ and a buyer by $b \in F$. A firm b in industry $j \in F$ 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-4 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.

5.2 *Deriving the estimating equation for the firm-level elasticity of substitution across industries*

In this section, we further derive conditions from the model to estimate the firm-level elasticity of substitution across industries. We now rewrite the maximization problem of the firm such that it maximizes

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

subject to (3) , (4), and $p_{i,bj} = \left(\sum_s \mu_{si,bj}p_{si,bj}^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}$. ζ is the firm-level elasticity of substitution across industries i we estimate. Notice that in this case, we need values for ϵ and $\mu_{si,bj}$. We consider $\epsilon = \hat{\epsilon}$, where $\hat{\epsilon}$ is our estimate, and we recover $\mu_{si,bj}$. All details about how we do this, and how we derive the expression are in Appendix C.2. The maximization problem yields the following expression:

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

where $p_{bj} = \left(\sum_{i'}\left(p_{i',bj}^{1-\zeta}w_{i',bj}\right)\right)^{\frac{1}{1-\zeta}}$ is a CES price index, $PM_{i,bj} \equiv p_{i,bj}x_{i,bj}$, and $PM_{bj} \equiv \sum_i PM_{i,bj}$, and $\log\left(w_{i,bj}\right)$ is the error term. This is our main estimating equation for the firm-level elasticity of substitution parameter ζ 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 $\widehat{\theta} = (\alpha, \zeta, \epsilon, \delta)$. We set the elasticity of substitution between different consumption varieties $\sigma = 4$, the elasticity of substitution between labor and the composite intermediate input $\alpha = 0.5$. We estimate the firm-level elasticity of substitution across suppliers (η) and the firm-level elasticity firm-level elasticity of substitution across industries (ζ) in the following section.

6.1 Addressing unobservable productivity shocks and entry/exit of suppliers

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:

$$\log\left(\frac{\widehat{PM}_{si,bj,t}}{\widehat{PM}_{i,bj,t}}\right) = (1-\epsilon)\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right) + \omega_{d(b),t} + \omega_{i,t} + \omega_{o(s)} + X\beta + \epsilon_{si,bj,t}. \tag{11}$$

For more details about how we arrive to this expression, see Appendix C. $\omega_{d(b),t}$ are destination district/month fixed effects, $\omega_{i,t}$ are industry/month fixed effects, 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 $\widehat{x}_t \equiv \frac{x_t}{x_{t-1}}$ to express variables in changes with respect to the previous month. Under the assumption that the geometric mean of productivity shocks across common suppliers of buyer b are time-invariant, we can then rewrite the main right-hand side variable as $\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right)$. In this variable, notice that the denominator is now $\widetilde{p}_{si,bj,t}$ instead of $p_{si,bj,t}$, where

$$\widetilde{p}_{i,bj,t} = \prod_{s\in\Omega^*_{i,bj,t}} p_{si,bj,t}^{\frac{1}{N^*_{i,bj,t}}}$$

is a geometric mean of unit values across common suppliers (i.e. suppliers for buyer b that appear in both current and previous month). 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 include controls X described in Section 3 and correction term following Feenstra (1994) to take into account the fact that sellers enter and exit in the data, which could induce a bias in the estimation of the elasticities. More details about this are in Appendix C.1.3. We include 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, depending on the nature of the shock. The logic behind clustering at the state level is that most policies, even those announced by the center, are implemented at the state level.

Now, from Equation (10), we arrive to the following expression we take directly to the data:

$$\log\left(\frac{\widehat{PM}_{i,bj,t}}{\widehat{PM}_{bj,t}}\right) = (1-\zeta)\log\left(\frac{\widehat{p}_{i,bj,t}}{\widehat{\widehat{p}}_{bj,t}}\right) + \omega_{d(b),t} + \omega_{i,t} + \epsilon_{i,bj,t}. \tag{12}$$

More details about this are in Appendix C.2

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 and in changes, this expression is

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

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 (13) below we formalize this intuition.

$$log(\widehat{p}_{si,bj,t}) = \beta^R Red_{o_{(s)}} Lock_t + \beta^O Orange_{o_{(s)}} Lock_t + \omega_{d_{(b)},t} + \omega_{i,t} + \omega_{o_{(s)}} + X\beta + \epsilon^{\nu}_{si,bj,t}. \tag{13}$$

$Lock_t$ is a dummy variable that equals 1 for the months of March and April 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.

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

$$log(\widehat{p}_{si,bj,t}) = \beta^R Red_{o_{(s)}d_{(b)}} Lock_t + \beta^O Orange_{o_{(s)}d_{(b)}} Lock_t + \omega_{d_{(b)},t} + \omega_{i,t} + \omega_{o_{(s)}} + X\beta + \epsilon^{\nu}_{si,bj,t}. \quad (14)$$

Details on how we obtain this functional form are contained in Appendix C.1.4. $Lock_t$ is a dummy variable that equals 1 for the months of March and April 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. Finally, we consider the average of the instruments across s to obtain unbiased estimators of ζ. More details are in Appendix C.2.3.

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 3 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 9a and 9b, 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 of both firm-level elasticities of substitution across suppliers and industries.

6.3.1. Firm-level elasticities of substitution across suppliers

First, we report OLS estimates in Table 5. The implied elasticities exhibit a robust value of 0.83 across all the different specifications. Our preferred specification is column (5), where we include product-month FEs, destination-month FEs, and origin FEs. To test whether our estimates vary widely depending on the level of aggregation, in column (6) we repeat the estimation for suppliers within the same HS-2 industry, and find that the estimates do not change. 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 6. We find evidence that inputs across different suppliers of a firm within the same HS-4 industry are highly complementary, ranging from 0.36-0.58, depending on the specification and the instrument we use. Our preferred estimate is 0.58, as reported in column (1), where we use both the seller and the seller-buyer level instrument, essentially exploiting variations arising from both seller level production costs and seller-buyer level transportation costs. This estimation also accounts for the entry and exit of suppliers through controlling for the Feenstra correction term. Not controlling for entry and exit of suppliers using the Feenstra term reduces the estimate to 0.36 in column (2). In columns (3) and (4), we repeat the estimation by only using variations in seller level production costs driven by the lockdown. As expected the F-stats are slightly lower than when we use both the instruments jointly, but still significant.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.

In table 7 we subject our main specification to three types of robustness checks: First, we change the period of lockdown dummy from March-April in our main specification to March-May to reflect the fact that in some districts the lockdown zones remained effective until May. The estimate slightly rises to .64, but remains within the confidence interval. Second, in column (2) we change the level of aggregation to define the industry in terms of the 2-digit HS code. This reduces the complimentarity, and the estimate increases to .72. Following Redding and Weinstein (2020) we have assumed that productivity shocks across common suppliers of a firm does not change over time. In the main estimation, we have defined common suppliers as firms who have also supplied inputs to the buyer in the last quarter (3 months). In column (3), to address the concern that entry and exit could be still affecting our estimates after controlling for the Feenstra term, we define the set of common suppliers to be those that supply to the firm in the last 2 months. Our estimate remains virtually unchanged. In column (4), when we consider the set of common suppliers to be those that supply to the firm in the last month. The estimates drop to .34, well within the conidence interval of the original estimate. These robustness checks show that our estimates are stable across various specifications, including correcting for entry and exit, changing the assumption on productivity shocks, and using two different instruments. In the following section, we use the estimated value of this elasticity across different suppliers of the same HS-4 product at 0.58, to analyze how shocks to small sets of firms can propagate through supply chains in the presence of complementarities across suppliers.

In table 8 and figure 8 we show the estimates of this elasticity of substitution across twenty broad industries. We find that classical measurement errors in unit values tend to produce very similar estimates of the elasticity of substitution across industries, with most values lying in the range .80-.90.[12] However, once we instrument for the unit values with the Covid-19 induced lockdown variations, we find that there is wide heterogeneity across industries in the estimate of this elasticity of substitution. While for the majority of the industries we find evidence for complementarity, there are some industries such as Plastics, Vegetables, and Handicrafts where suppliers across substitutes.

6.3.2. Firm-level elasticities of substitution across industries

In table 9 we report the results of the OLS estimation of the firm-level elasticity of substitution across industries. Across various combinations of fixed effects, we find a very robust value of this elasticity of substitution at 0.87, suggesting again complementarity in the usage of inputs. Consistent with our elasticity of substitution across suppliers, we report the results of this estimation at HS-2 level in column (5) and find that the estimate is virtually unchanged. For similar reasons as discussed in the previous sections, it is very likely that our estimators are contaminated by simultaneous demand shocks that happened during Covid-19. In table 10 we report the results of the IV estimation. Column (3) reports the results from our main estimation where the treatment period is March-April 2020 and the common suppliers are defined over the last quarter. The elasticity of substitution across industries is 0.57. In columns (1) and (2) we show that this estimate varies between .51 and .65, depending on how common suppliers are defined for the estimation of the elasticity of substitution across suppliers. This estimate is more sensitive to how the treatment period is defined and increases to .84 when the treatment period is defined between March-May. However, across all specifications, we find that inputs are highly complementary. In the literature, people have estimated a wide range of values for

[12] In the trade literature it is widely recognized that deriving unit values by dividing trade values with quantities produces some measurement error.

the elasticity of substitution across industries, depending on the aggregation of the industry and the question under study. Our numbers are very close to Boehm et al. (2019) who estimate this elasticity to be 0.40-0.62 for different inputs at HS-10 for non-Japanese firms and 0.2 for Japanese firms. Atalay (2017) finds this estimate to be around .1 for 30 aggregated industries using the US data. In our final simulations exercise in future versions, we will consider a wide range for this elasticity of substitution, as suggested by our estimates.

6.4 *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. 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 4-digit HSN industries, (ii) firm sales to 4-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 firm to firm trade data from March 2019 to February 2020, that is, the pre-Covid period. Labor employed and final sales by firms are obtained by merging our firm level data with data from Indiamart, which contains information on firm-level employment and final sales.[13] For more details for this, see Appendix A.

There are 1293 industries. The average firm is connected to 10 industries as a buyer and 5 industries as a seller. The most connected buyer and seller buys from and sells to over 500 industries. We use this 94,555 by 94,555 input output matrix consisting of firm-level sales and purchases from these 1293 industries at the HS-4 level to understand how complemenatarities at the firm-level affect the propagation of shocks through the firm production networks.

6.4.1. How much does the firm-level elasticity of substitution matter?

29% of all firms lie in the red zone. In this counterfactual, we shock the productivity of all firms located in the red zone by 15%. We find that a 15% productivity shock to firms in the red zone reduces GDP by 6.05%. This fall would be 0.45 pp less in a model where firms in the same HS-4 industry are considered substitutes ($\epsilon = 1.75$), and 0.22 pp more when firms in the same HS-4 industry are considered almost leontieff ($\epsilon = .001$) . In terms of GDP losses, given that the quarterly GDP of this state was close to 32.5 billion USD in 2020-2021, the additional losses due to firm-level complementarities translate into 146.2 million USD (4.2 USD per capita), compared to the case when firms are substitutes. To put these numbers into perspective, Baqaee and Farhi (2019) showed that complementarities at the industry level (with an elasticity of substitution .001) amplify the effect of a negative 13% shock in the oil-industry

[13] `https://www.indiamart.com/`

on GDP by about 0.61%.

How important are these second order effects that we have estimated? To assess the importance of these second order effects, we simulate different levels of negative productivity shocks for 4 different values of elasticities of substitution and plot the second order percentage point change in GDP due to these shocks in figure 11. The top two plots show these differences for a very high level of complementarity .001 and our estimated elasticity .58, respectively. The bottom two pictures show the additional change in GDP due to the second order when firms are substitutes. Two things are clear from these pictures: First, for the same negative productivity shock, the second order effects are much larger when the firms are more complements. Second, given the same value of the elasticity of substitution, as the magnitudes of the productivity shocks increase, the second order effects become more and more important, that is, there are non-linear effects. Observing the bottom two pictures, we can clearly see that as firms become more and more substitutes, the second order effects actually dampen the negative first order effects, and more so, for higher values of productivity shocks. That is, unlike the first order effects which only depend on firm sizes, complementarities at the firm-level non-linearly amplify the effects of negative productivity shocks.

These graphs illustrate to us the importance of second order effects that are largely driven by complementarities at the firm level, especially for large short-lived negative productivity shocks such as Covid-19. For a very long time, since Hulten's 1978 theorem, policy makers and researchers have emphasized the importance of firm sizes in the propagation of shocks. In the next counterfactual we are going to investigate how important are large firms versus connected firms in the propagation of shocks.

6.4.2. Counterfactual GDP changes for large and connected firms

In this counterfactual, we look at the fall in GDP if the largest or the most connected firms in the red zone are allowed to operate. Largest firms are measured by domar weights (final sales share). The most connected firms are measured by the Leontief inverse which measures the direct and indirect connections of suppliers. In figure 12, we plot the results from conducting two different counterfactuals for two different values of elasticities of substitution: .001 and .58, where in the first counterfactual the largest 10% firms in the red zones are allowed to operate and in the second counterfactual, the most connected 10% firms in the red zones are allowed to operate. In the x-axis, we plot different values of negative productivity shock, starting from -5 to -35. In the y-axis, the blue graph represents the percentage point difference in GDP when the largest 10% firms are allowed to operate compared to the baseline case and likewise, the red graph represents the same thing but when the most connected firms are allowed to operate. Two things are notable from these graphs: First, as the level of complementarity increases, it pays more to save the more connected firms. Second, for the same level of complementarity, as the magnitude of the shock increases, it pays more to save the more connected firms too. This counterfactual illustrates that when evaluating which firms to value more when it comes to the affects on aggregate GDP, policy makers should be looking at not just large firms but also the connected firms. In fact, the connected firms become more important for a large yet short-lived shock and when suppliers are highly complementary.

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, can explain part of the GDP decline in India observed during the Covid-19 crisis of 2020. Our analysis shows that these complementarities matter more for large negative productivity shocks. As the magnitude of the shock increases, so does the importance of the most connected firms in the the propagation of shocks.

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.

Bas, M., A. M. Fernandes, and C. Paunov (2022). How Resilient Was Trade to COVID-19.

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.

Chakrabati, A., K. Mahajan, and S. Tomar (2021). Aggregate shocks, domestic trade collapse and regional realignment.

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. (1994). New product varieties and the measurement of international prices. *The American Economic Review*, 157–177.

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.

Taschereau-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: Sales shares, by HSN section

HSN section	Share
Vegetables	13.0233
Machinery	12.5682
Metal	10.7311
Plastics	10.3752
Minerals	10.0090
Chemicals	9.5806
Transport equipment	7.2062
Textiles	6.2959
Processed foods	5.1820
Fats	2.3189
Wood	1.8621
Jewelry	1.8327
Surgical instruments	1.7636
Handcrafts	1.6757
Wood derivatives	1.3398
Miscellaneous	1.3318
Animals	1.2548
Clothing	1.0617
Art	.4029
Leather	.1711
Arms and ammo	.0126

Notes: In the first column we report the name of the HSN section. In the second column we calculate the sales share by HSN section.

TABLE 3: Sales shares, by type of transaction

HSN section	Share
Within-state	72.6822
Inter-state	23.2183
Exports	4.0994

Notes: For all sellers located within our Indian state, we calculate sales shares where the destination was either within the state, to another state of India, or exports.

TABLE 4: Purchase shares, by type of transaction

HSN section	Share
Within-state	52.2224
Inter-state	44.5151
Imports	3.2623

Notes: For all buyers located within our Indian state, we calculate sales shares where the destination was either from within the state, from another state of India, or imports.

Table 5: OLS, firm-level elasticity of substitution across suppliers

	(1)	(2)	(3)	(4)	(5)	(6)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.1691	0.1699	0.1693	0.1692	0.1700	0.1621
	(0.0111)	(0.0112)	(0.0112)	(0.0112)	(0.0113)	(0.0081)
R^2	0.3896	0.3916	0.3901	0.3897	0.3920	0.3716
Obs	3175672	3173948	3174893	3175653	3173147	2769880
ϵ	0.8308	0.8300	0.8306	0.8307	0.8299	0.8378
HSN $\times$ Month FE		Y			Y	Y
Destination $\times$ Month FE			Y		Y	Y
Origin FE				Y	Y	Y

Notes: These are the estimators resulting from running OLS regressions from equation (11). The first row reports the estimator associated to changes in relative unit values in logs. 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 six columns. In columns (1)-(5), an industry is 4-digit HSN codes. In column (6), an industry is 2-digit HSN codes. Each column is comprised by a different set of fixed effects. The last three rows denote which fixed effects are included. All specifications include the controls mentioned in the paper.

Table 6: 2SLS, firm-level elasticity of substitution across suppliers

	(1)	(2)	(3)	(4)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.4190	0.6358	0.5057	0.4493
	(0.2457)	(0.2528)	(0.2573)	(0.3721)
Obs	2619498	2619498	2619498	2619498
K-PF	57.428	58.380	15.194	15.221
J-stat	3.638	3.989	2.999	4.047
ϵ	0.5809	0.3641	0.4942	0.5506
Seller/Buyer IV	Y	Y		
Seller IV	Y	Y	Y	Y
Feenstra term	Y		Y	

Notes: These are the estimators resulting from running 2SLS regressions from equation (11). The set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \left(\Omega_{i,bj,t-1} \cup \Omega_{i,bj,t-2} \cup \Omega_{i,bj,t-3}\right)$. That is, a supplier s of buyer b is considered *common* if they traded during the two previous months. The first stage uses either seller/buyer or seller-level instruments. Seller/buyer-level instruments correspond to equation (14), while seller-level instruments correspond to equation (13). The first row reports the estimator associated to changes in relative unit values in logs. The second row reports standard errors in parentheses. The fourth row reports the Kleibergen-Paap F statistic from the first stage. The fifth row reports Hansen J statistic. The sixth row reports the implied value for ϵ, which is 1 minus the estimator on the first row. An industry is 4-digit HSN codes and the treatment period is March-April 2020. The table contains four columns. Each column corresponds to different combinations of instruments, and whether we are accounting for the Feenstra correction term as described in C.1.3. The last three rows indicate which set of instruments we are using or whether we are including the Feenstra correction term. Standard errors for columns (1)-(2) are clustered at the origin/destination state level, while columns (3)-(4) are clustered at the origin state level. All specifications include the controls mentioned in the paper.

TABLE 7: 2SLS, firm-level elasticity of substitution across suppliers, robustness

	(1)	(2)	(3)	(4)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.3576	0.2743	0.4539	0.6552
	(0.2400)	(0.2095)	(0.3071)	(0.3077)
Obs	2619498	2257290	2537949	2294206
K-PF	65.244	9.149	20.566	67.743
J-stat	6.187	4.383	4.594	3.382
ϵ	0.6423	0.7256	0.5460	0.3447
Spec.	Mar-May	2-digit HSN	2 prev. months	1 prev. months

Notes: These are the estimators resulting from running 2SLS regressions from equation (11). The first stage uses both seller/buyer and seller-level instruments. Seller/buyer-level instruments correspond to equation (14), while seller-level instruments correspond to equation (13). The first row reports the estimator associated to changes in relative unit values in logs. The second row reports standard errors in parentheses. The fourth row reports the Kleibergen-Paap F statistic from the first stage. The fifth row reports Hansen J statistic. The sixth row reports the implied value for ϵ, which is 1 minus the estimator on the first row. The table contains four columns. Each column introduces changes with respect to our main specification, which is described in the last row. In column (1), the treatment period is March-May 2020. In column (2), an industry is 2-digit HSN codes. In column (3), the set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \left(\Omega_{i,bj,t-1} \cup \Omega_{i,bj,t-2}\right)$. In column (4), the set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \left(\Omega_{i,bj,t-1}\right)$. Standard errors are clustered at the origin/destination state level. All specifications include the controls mentioned in the paper.

TABLE 8: Firm-level elasticities of substitution across suppliers, by HSN section

Section	Name	OLS elast.	2SLS elast.
1	Animals	0.8407	0.4525
2	Vegetables	0.8823	1.0557
3	Fats	0.8827	0.9548
4	Processed foods	0.8505	0.4083
5	Minerals	0.9147	0.8947
6	Chemicals	0.8536	0.5289
7	Plastics	0.8613	1.0897
8	Leather	.	.
9	Wood	0.9361	0.6464
10	Wood derivatives	0.8960	0.8926
11	Textiles	0.8920	0.5541
12	Clothing	0.8984	0.5791
13	Handcrafts	0.8232	1.1550
14	Jewelry	0.8578	0.3527
15	Metal	0.9101	0.8176
16	Machinery	0.7141	0.6964
17	Transport equipment	0.6332	0.4846
18	Surgical instruments	0.7446	0.8326
19	Arms and ammo	0.5132	.
20	Miscellaneous	0.8272	1.1696
21	Art	.	.

Notes: Each row corresponds to an industry, which is defined as a HSN section. The second column contains the name of the industry. The third and fourth columns report the implied elasticity by OLS and 2SLS as in Equation (11). Both OLS and 2SLS estimators include HSN/month, destination/month, and origin fixed effects, and standard errors are clustered at the origin/destination state level. All specifications include the controls mentioned in the paper. Missing elasticities were not able to be estimated due to lack of statistical power.

TABLE 9: OLS, firm-level elasticity of substitution across industries

	(1)	(2)	(3)	(4)	(5)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.1314	0.1296	0.1314	0.1295	0.1186
	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0040)
R^2	0.2164	0.2340	0.2169	0.2343	0.2056
Obs	1824860	1822979	1823996	1822112	1345656
ζ	0.8685	0.8703	0.8685	0.8704	0.8813
HSN $\times$ Month FE		Y		Y	Y
Destination $\times$ Month FE			Y	Y	Y

Notes: These are the estimators resulting from running OLS regressions from equation (12). The first row reports the estimator associated to changes in relative unit values in logs. Standard errors are clustered at the destination state level, and are reported in parentheses below each estimator. The fifth row reports the implied value for ζ, which is 1 minus the estimator on the first row. The table contains five columns. In columns (1)-(4), an industry is 4-digit HSN codes. In column (5), an industry is 2-digit HSN codes. Each column is comprised by a different set of fixed effects. The last three rows denote which fixed effects are included.

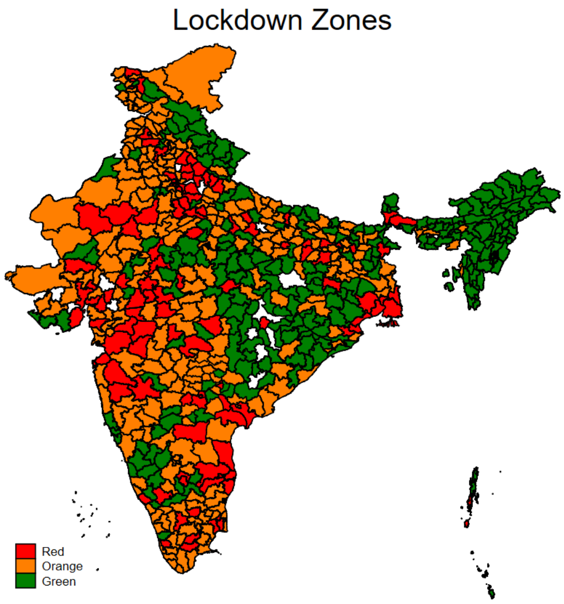
TABLE 10: 2SLS, firm-level elasticity of substitution across industries

	(1)	(2)	(3)	(4)	(5)
$\log\left(\frac{\hat{p}}{\hat{\bar{p}}}\right)$	0.4866	0.3469	0.4250	0.1585	0.3132
	(0.6185)	(0.4231)	(0.4668)	(0.2556)	(0.6934)
Obs	1822112	2058653	2141783	1345656	1822112
K-PF	15.264	33.158	18.760	14.355	12.993
J-stat	3.4055	1.9786	2.5250	11.4321	12.0848
ζ	.5133	.6530	.5749	.8414	.6867
Number of prev. months of common suppliers	1	2	3	1	1
Specs.				Mar-May	2-digit HSN

Notes: These are the estimators resulting from running 2SLS regressions from equation (12). The set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$. That is, a supplier s of buyer b is considered *common* if they traded during the two previous months. The first stage uses either seller/buyer or seller-level instruments. Seller/buyer-level instruments correspond to equation (14), while seller-level instruments correspond to equation (13). The first row reports the estimator associated to changes in relative unit values in logs. The second row reports standard errors in parentheses. The fourth row reports the Kleibergen-Paap F statistic from the first stage. The fifth row reports Hansen J statistic. The sixth row reports the implied value for ϵ, which is 1 minus the estimator on the first row. The seventh row indicates which set of common suppliers we consider. A value of 1 implies that common suppliers appear in both current and previous month; for a value of 2, in current and two previous months; for a value of 3, in current and three previous months. The last row indicates additional specifications for columns (4) and (5). The table contains five columns. In columns (1)-(3), an industry is 4-digit HSN codes and the treatment period is March-April 2020. In column (4), an industry is 4-digit HSN codes and the treatment period is March-May 2020. In column (5), an industry is 2-digit HSN codes and the treatment period is March-April 2020. Standard errors are clustered at the origin/destination state level.

FIGURES

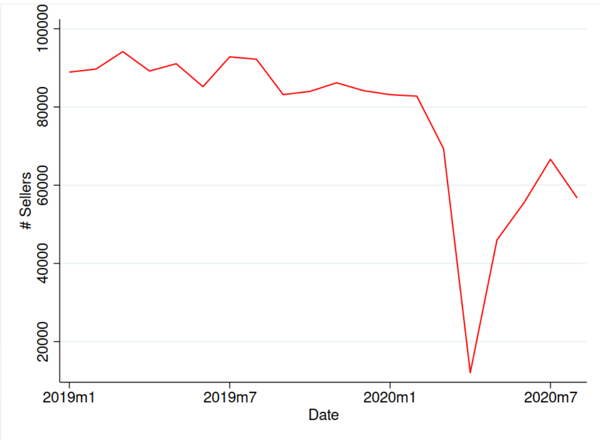
FIGURE 1: MAP SHOWING INDIA'S LOCKDOWN ZONES



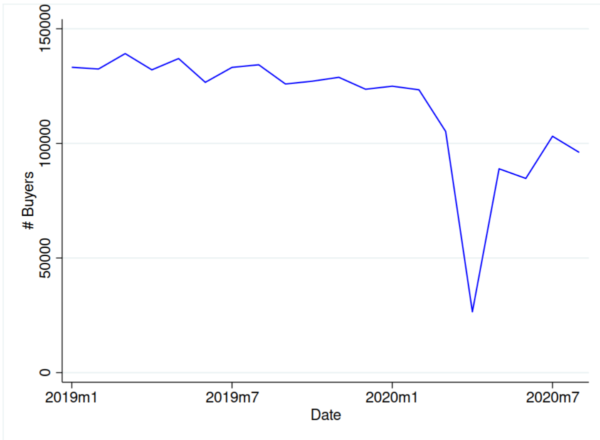
Notes: The map above shows the lockdown zones across Indian districts, where the lockdown was announced on March 25,2020.

FIGURE 2: VARIATION OVER TIME

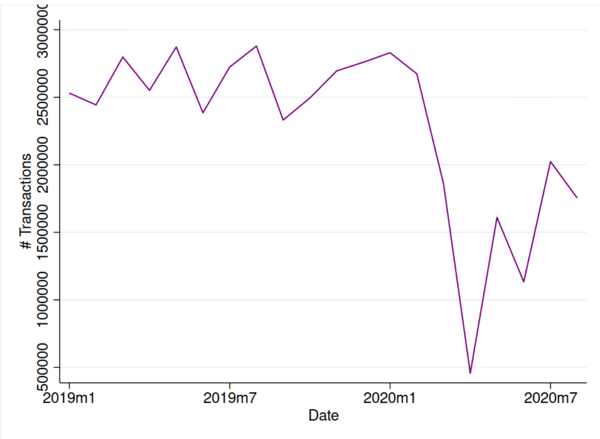
(a) Number of sellers



(b) Number of buyers



(c) Number of transactions

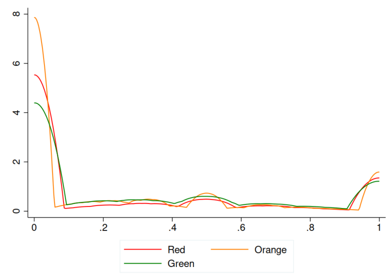


(d) Total sales

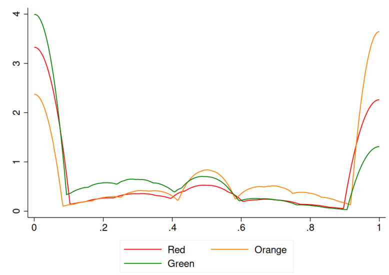
Notes: This figure is comprised by 4 panels. In each panel, the horizontal axis is a month, and the vertical axis is determined by the panel. In the first panel, we show the number of sellers that reported a transaction by month. In the second panel, we show the number of buyers that reported a transaction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

FIGURE 3: SHARE DISTRIBUTIONS OF COLORS

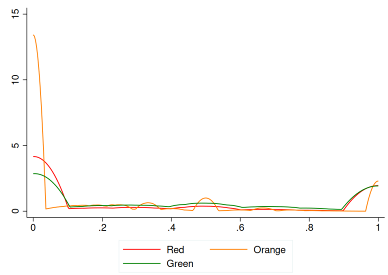
(a) Sellers in Red



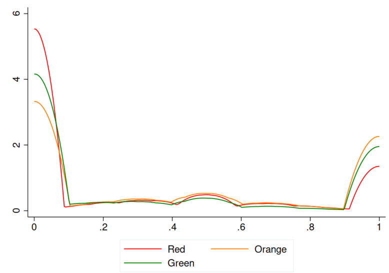
(b) Sellers in Orange



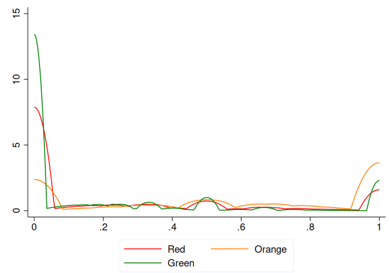
(c) Sellers in Green



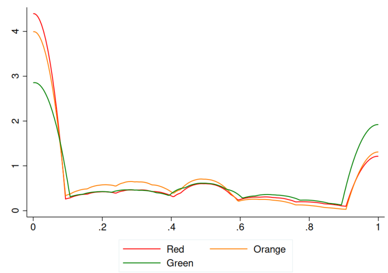
(d) Buyers in Red



(e) Buyers in Orange

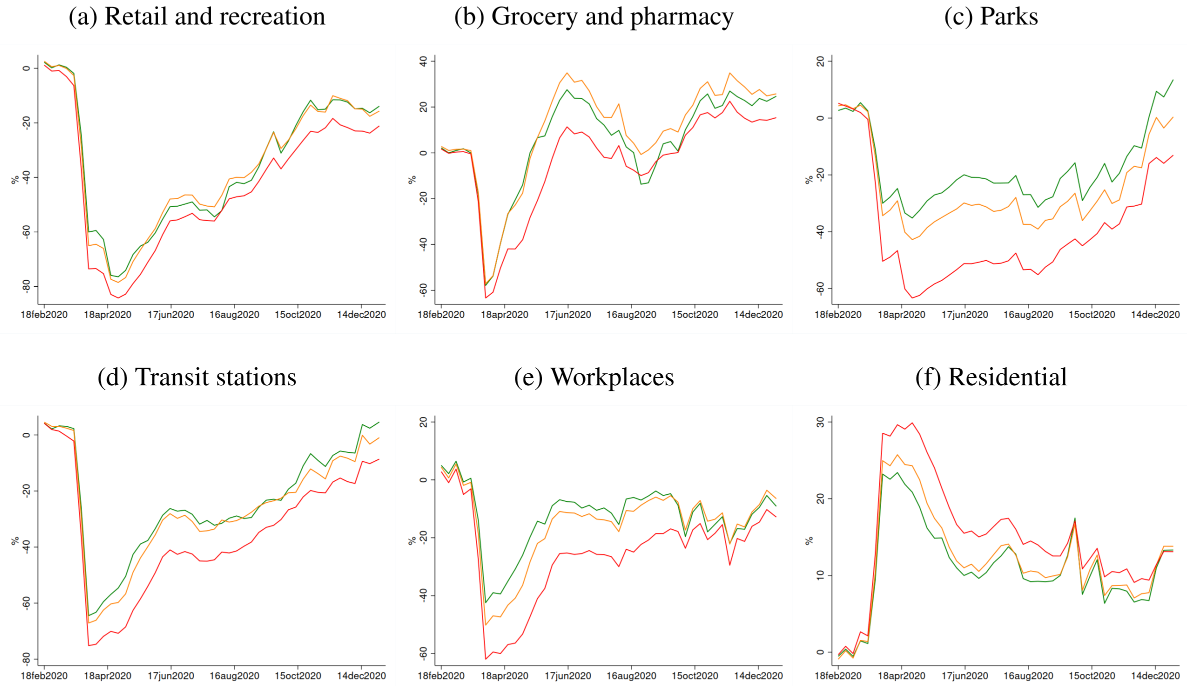


(f) Buyers in Green



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.

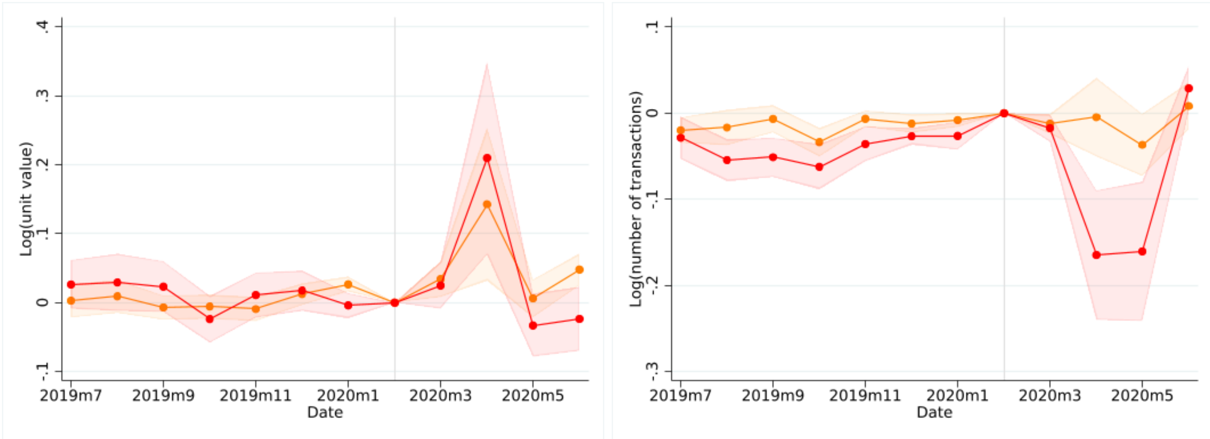
FIGURE 4: GOOGLE MOBILITY TRENDS



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

FIGURE 5: 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.

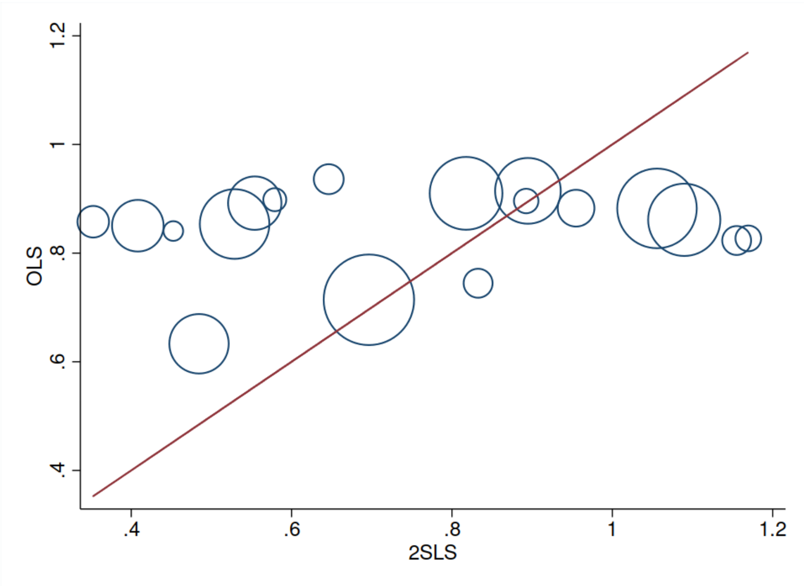
FIGURE 6: UNIT VALUE, SELLER-BUYER REGRESSIONS

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 industry/month, origin district, and destination district fixed effects, and standard errors are clustered at the origin/destination state level. An industry is 4-digit HSN codes. All controls mentioned in the paper are included. The vertical line in January 2020 splits pre and post periods. The baseline category are sellers and buyers located in *Green* districts on January 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 7: 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 industry/month, origin district, and destination district fixed effects, and standard errors are clustered at the origin/destination state level. An industry is 4-digit HSN codes. All controls mentioned in the paper are included. The vertical line in January 2020 splits pre and post periods. The baseline category are sellers and buyers located in *Green* districts on January 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 8: ELASTICITIES BY SELLER INDUSTRY

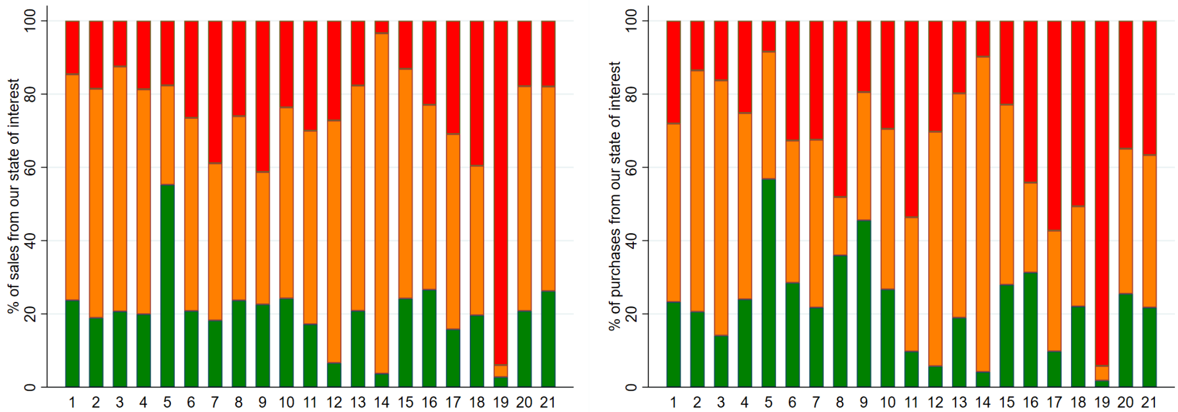


Notes: The vertical axis is the firm-level elasticity of substitution by the industry of the seller, estimated by OLS. The horizontal axis is estimated by 2SLS. An industry is an HSN section. The size of each bubble is determined by total sales in the corresponding industry.

FIGURE 9: % OF SALES/PURCHASES, BY COLOR OF DESTINATION DISTRICTS

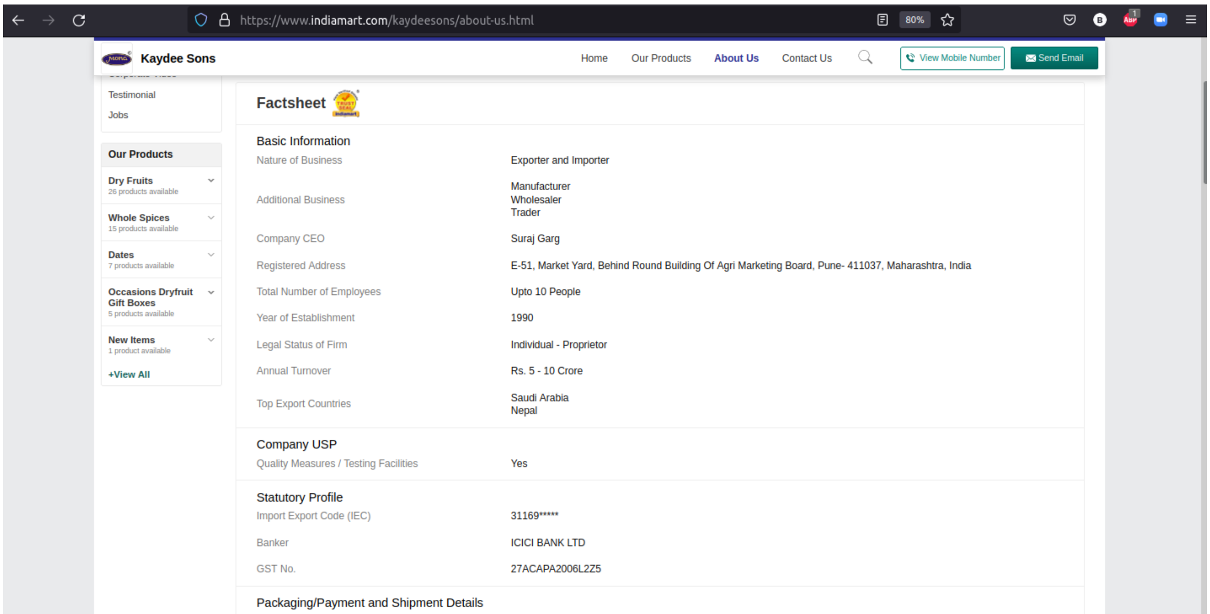
(a) Sales

(b) Purchases



Notes: On the left (right) panel, for each HSN section (horizontal axis), we plot the share of total sales (purchases) of firms located in our large Indian state by color of selling (buying) districts.

FIGURE 10: INDIAMART



Notes: This vignette shows an example on how a firm in the website IndiaMART looks like. This is the factsheet firm called Kaydee Sons that sells dry fruits, dates, spices, seeds, walnut kernels, gift pack, and fresh apricot. We scrape all the variables included in this vignette, where the GST number, number of employees, and annual turnover are the key variables.

FIGURE 11: HOW IMPORTANT ARE SECOND ORDER EFFECTS?

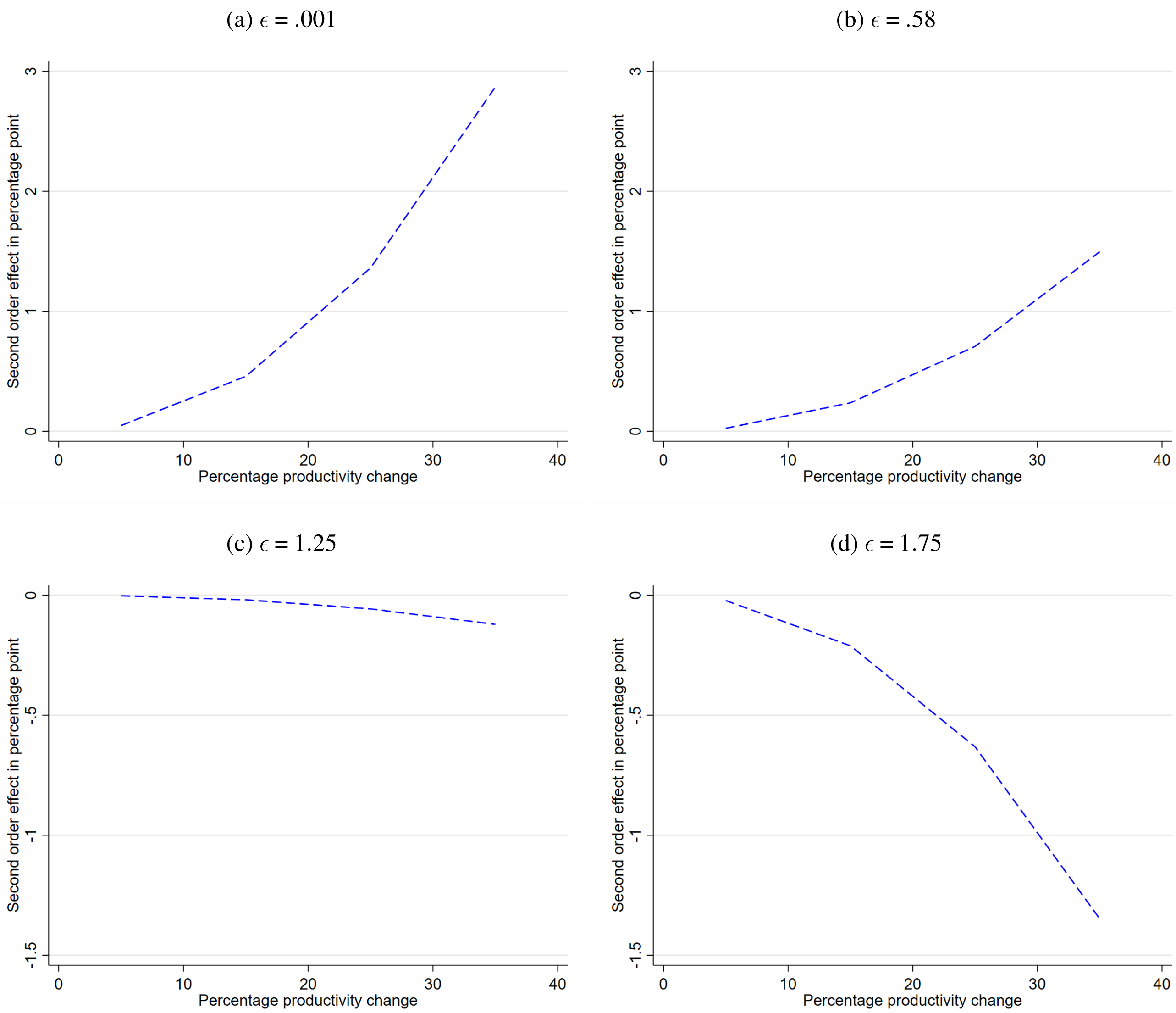
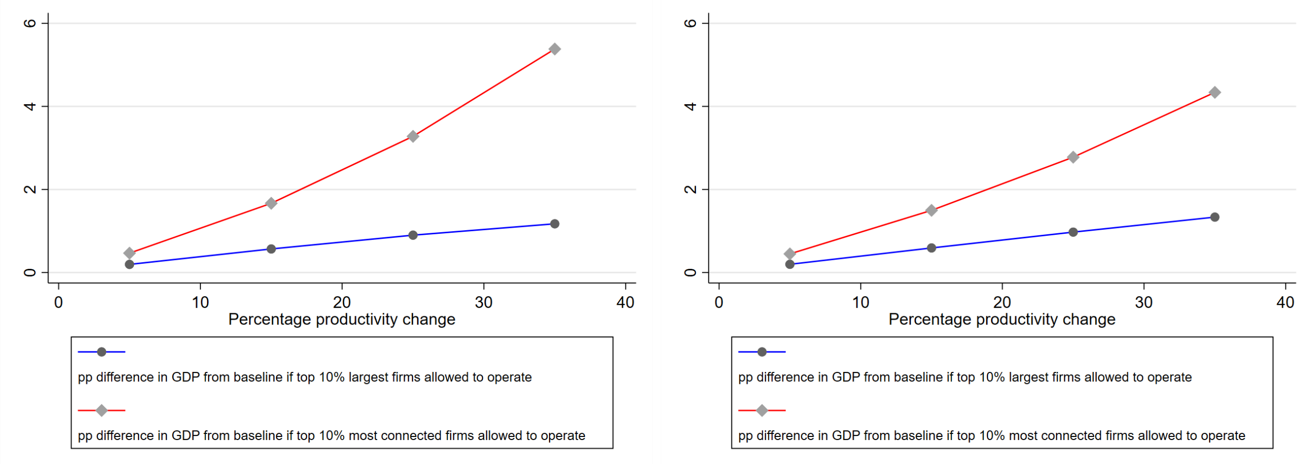


FIGURE 12: LARGE OR CONNECTED FIRMS?

(a) $\epsilon = .001$ (b) $\epsilon = .58$



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

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

Labor and sales. Our firm-to-firm dataset lacks data on number of employees and final sales. Then, the objective is to predict values for number of employees and final sales for all buyers and sellers of the dataset. We do this by obtaining data on number of employees and total sales from an external dataset for a subset of our firms, run an OLS regression of both labor and final sales on observable variables in our firm-to-firm dataset, store the elasticities, and use them to predict labor and final sales for all our firms.

We scraped the website *IndiaMART*, which is the largest B2B digital platform in India. The website contains subpages with a profile per vendor. In Figure 10 there is an example of vendor that sells dry fruits, spices, among others. We scraped around around 300-400K firm profiles, and then sent them to the tax authority to be matched with our firm-to-firm trade dataset. The matching procedure yielded 50,720 unique firms.

Each firm reports its number of employees and annual turnover (sales), both reported in brackets. The reported brackets for sales are: up to 50 Lakh, 50 Lakh-1 Crore, 1-2 Crore, 2-5 Crore, 5-10 Crore, 10-25 Crore, 25-50 Crore, 50-100 Crore, 100-500 Crore 500-1,000 Crore, 1,000-5,000 Crore, 5,000-10,000 Crore, more than 10,000 Crore. First, we convert each reported number into rupees, since sales in the trade dataset is reported in rupees.[14] Then, for each firm we assign the median value of its corresponding sales bracket. For the last bracket, we consider the upper bound to be 100,000 Crore. The reported brackets for labor are: up to 10 employees, 11-25, 26-50, 51-100, 101-500, 501-1000, 1001-2000, 2001-5000, more than 5000 employees. For each firm we assign the median value of its corresponding labor bracket. For the last bracket, we consider the upper bound to be 50,000 employees.

We run the following OLS regressions:

$$\begin{aligned}\log(labor_n) &= \alpha_0 + \alpha_1 \log(sales_n) + \alpha_2 \log(distance_n) + \epsilon_i^l \\ \log(final_n) &= \beta_0 + \beta_1 \log(sales_n) + \beta_2 \log(distance_n) + \epsilon_i^f,\end{aligned}$$

where $sales_n$ are total sales of intermediates of firm n and $distance_n$ is the average distance in KM of all transactions reported, $labor_n$ is the number of employees constructed as previously explained, and $final_n$ is final sales. We constructed final sales by subtracting intermediate sales from total sales. We constructed total as we previously explained, and we calculate intermediate sales from our firm-to-firm dataset. In the vast majority of cases, this difference was positive, which reassures that IndiaMART reports total sales. Whenever the differences were negative, we input a value of 0 instead, which means that all sales of that firm are for intermediates.

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

B Dijkstra algorithm

In this section we list the steps of a Dijkstra algorithm we used to construct our the seller/buyer level instruments. The steps are the following:

1. First, we obtained a set of shapefiles of district administrative boundaries for India according to India's 2011 census;

2. We then reprojected the shapefiles into an *Asian/South Equidistance Conic* projection, which is the projection that preserves the notion of distance the best;

3. Once shapefiles are reprojected, we obtain the centroid of each district in India;

4. We now have to construct a network structure according to the set of centroids. There are many ways to construct a network, so we need to take a stance on how to form the connections between nodes (i.e. centroids). For each centroid, we generate connections to the k closest centroids according to Euclidean distances.[15] We follow Fajgelbaum and

[14] 100,000 rupees = 1 Lakh; and 10,000,000 rupees = 1 Crore.

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

Schaal (2020) and consider $k = 8$ such that we consider the main cardinal directions (i.e. north, south, east, west, north-east, south-east, north-west, south-west). This is a network structure we can now feed into a 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 the district pair, and the distance of 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, and obtain our seller/buyer level instruments. Our first instrument is the share of districts in a route that are *Red*, *Orange*, or *Green*. When calculating these shares, we rule out the zone where the buyer resides so we don't consider demand-side shocks in our instrument. Using the distance of each leg, our second instrument is the share of meters of the route that are *Red*, *Orange*, or *Green*. We consider a leg to be of color $x = Red, Orange, Green$ whenever the origin district was of color x. In this case we also ignore the color of the district where the buyer resides.

C DERIVATIONS

C.1 *Estimation of firm-level elasticities of substitution across suppliers*

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

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

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

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

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

where $(\ldots)$ are components that we do not write in detail since they will cancel as we progress in the derivations. 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.1.2. Addressing unobservable productivity shocks

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

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

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

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

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

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

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

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

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

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

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

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

This expression in changes is

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

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

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

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

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

Then, the last term of our expression is

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

So our final expression boils down to

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

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

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

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

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

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

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

C.1.3. Addressing entry/exit of suppliers

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

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

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

$$s_{si,bj,t}\equiv\frac{PM_{si,bj,t}}{\sum_{s\in\Omega_{i,bj,t}}PM_{si,bj,t}},$$

$$\lambda_{i,bj,t}\equiv\frac{\sum_{s\in\Omega^*_{i,bj,t}}PM_{si,bj,t}}{\sum_{s\in\Omega_{i,bj,t}}PM_{si,bj,t}},$$

$$s^*_{si,bj,t}\equiv\frac{PM_{si,bj,t}}{\sum_{s\in\Omega^*_{i,bj,t}}PM_{si,bj,t}}.$$

In changes, the expression for expenditure shares is

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

Then, the geometric mean for expenditure shares is

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

So the final expression we take to the data is

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

C.1.4. Addressing endogeneity concerns

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

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

where $p_{si,t}$ is the marginal cost (MC) of production of good i for seller s in month t, $\tau_{sb,t}$ is the iceberg cost of transporting the good from seller s to buyer b in month t.

Now, we can then express this in changes, such that

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

In logarithms, we have

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

These two components of price inspire two sources of variation: The first is our seller level instrument, which utilizes variations in MC at the seller-product level due to lockdown measures at the seller's district. To isolate the variations in MC that is driven by the lockdown zones of the seller, we interact the lockdown dummy ($Lock_t$) which takes the value 1 in April and May with the lockdown zone (Red or Orange) of the seller.

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

Now we explain how we construct the instrument at the seller/buyer level. We have to take a stance about the functional form of the trade cost τ. We assume that trade costs are proportional to the travel time of the transportation of intermediate inputs, such that

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

If we express this in changes, we get

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

We exploit variation from the Covid 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. Combining the expression for changes in travel time due to the lockdown and trade costs, we get the following expression for our seller/buyer level instrument

$$
\begin{aligned}
log(\widehat{\tau}_{sb,t}) = \beta^R Red_{o_{(s)}d_{(b)}} Lock_t + \beta^O Orange_{o_{(s)}d_{(b)}} Lock_t \\
+ \omega_{d_{(b)},t} + \omega_{o_{(s)}} + X\beta + \epsilon^{\nu}_{si,bj,t}
\end{aligned}
$$

C.2 Estimation of firm-level elasticities of substitution across industries

In this section we describe steps we do to derive the firm-level elasticity of substitution across industries. First, we describe the model and the equations we take to the data. Second, we describe the additional inputs we need to take the equations to the data. In particular, we construct prices by leveraging our estimated elasticities of substitution across suppliers and the recovered productivity shocks. Finally, we describe the instrument we use to estimate this elasticity.

C.2.1. Expressions to estimate firm-level elasticities of substitution across industries

We rewrite the initial maximization problem, so

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

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

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

where $(\ldots)$ are components that we do not write explicitly since they will cancel out as we progress in the derivations. Now consider the same first order conditions with respect to $x_{i',bj}$ and divide them, such that

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

where $PM_{bj} \equiv \sum_i PM_{i,bj}$, and $p_{bj} = \left(\sum_i w_{i,bj}p_{i,bj}^{1-\zeta}\right)^{\frac{1}{1-\zeta}}$. As we did for the estimation of the elasticity of substitution across suppliers, we introduce a time dimension, apply Shephard's lemma to this CES price function, and assume that the geometric mean of productivity shocks across industries are constant, and obtain:

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

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

$$
\begin{aligned}
PM_{i,bj,t}\,p_{bj,t}^{1-\zeta} &= w_{i,bj,t}\,p_{i,bj,t}^{1-\zeta}\,PM_{bj,t},\\
\widehat{PM}_{i,bj,t}\,\widehat{p}_{bj,t}^{1-\zeta} &= \widehat{w}_{i,bj,t}\,\widehat{p}_{i,bj,t}^{1-\zeta}\,\widehat{PM}_{bj,t},\\
\log\left(\frac{\widehat{PM}_{i,bj,t}}{\widehat{PM}_{bj,t}}\right) &= (1-\zeta)\log\left(\frac{\widehat{p}_{i,bj,t}}{\widehat{p}_{bj,t}}\right)+\log\left(\widehat{w}_{i,bj,t}\right),\\
\log\left(\frac{\widehat{PM}_{i,bj,t}}{\widehat{PM}_{bj,t}}\right) &= (1-\zeta)\log\left(\frac{\widehat{p}_{i,bj,t}}{\frac{\widehat{\widetilde{p}}_{bj,t}}{\widehat{\widetilde{s}}_{bj,t}^{\frac{1}{1-\zeta}}}}\right)+\log\left(\widehat{w}_{i,bj,t}\right),\\
\log\left(\frac{\widehat{PM}_{i,bj,t}}{\widehat{PM}_{bj,t}}\right) &= (1-\zeta)\log\left(\frac{\widehat{p}_{i,bj,t}}{\widehat{\widetilde{p}}_{bj,t}}\right)+\log\left(\widehat{\widetilde{s}}_{bj,t}\right)+\log\left(\widehat{w}_{i,bj,t}\right).
\end{aligned}
$$

This is the expression we actually we use to estimate ζ. Notice that we have applied the time-invariance of productivity shocks $w_{i,bj,t}$ across all industries, not across common industries. This was necessary when estimating ϵ since each supplier is considered a variety, but not necessary in this case.

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

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

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

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

1. Run the 2SLS regression to obtain the estimator $\widehat{\epsilon}$;
2. Recover predicted values for the error term $\widehat{\phi}_{si,bj,t}$;
3. Calculate the empirical mean and variance of $\widehat{\phi}_{si,bj,t} : \left\{\widehat{\mu}_\phi, \widehat{\sigma}^2_\phi\right\}$;
4. Recover the values for mean and variance of $\log\left(\mu_{si,bj,t}\right), such that : (i)\mu = \widehat{\mu}_\phi and \sigma^2 = \frac{\widehat{\sigma}^2_\phi}{2}$;
5. Make a random draw for $\log\left(\mu_{si,bj,0}\right)$, which is drawn from a normal distribution with mean $\widehat{\mu}_\phi$ and variance $\frac{\widehat{\sigma}^2_\phi}{2}$;
6. For a given $\mu_{si,bj,0}$, recover $\mu_{si,bj,t}$ according to the following law of motion:

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

7. We then construct unit values by

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

C.2.3. Addressing endogeneity concerns

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

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

For intuition, consider the instrument that varies across both the color zone of the seller and the buyer (i.e. the share of districts of color red in the route between the location of the seller and of the buyer). Then, the new instrument is the simple average of these shares across sellers. Intuitively, the higher the shares of red-colored locations within the routes, the higher the shock on prices