Essays on International Economics

by

Brian Cevallos Fujiy

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2023

Doctoral Committee:

Professor Andrei Levchenko, Co-Chair
Associate Professor Sebastian Sotelo, Co-Chair
Professor Jagadeesh Sivadasan
Professor Dmitriy Stolyarov

Brian Cevallos Fujiy

bcfujiy@umich.edu

ORCID iD: 0000-0002-7973-7835

ACKNOWLEDGEMENTS

I thank my committee members, Andrei Levchenko, Sebastian Sotelo, Jagadeesh Sivadasan, and Dmitriy Stolyarov, who helped me tremendously to develop my dissertation. I also thank Dominick Bartelme who guided me through the initial stages of my job market paper. I specially thank my co-chairs and mentors, Andrei and Sebastian, for generously sharing their knowledge and experience with me. I thank my current co-authors for their hard work and patience during the job market. I specially thank Gaurav Khanna for his guidance and support during the Ph.D.

I thank my cube mates, Nadim, Joserra, and Sebastian, who provided me with their support and friendship. I thank Luis E., Luciana, and Hiroshi, for lodging whenever I visited Ann Arbor. I thank Luis A., Luis B., and Ben for their guidance during the Ph.D. I thank Rafael, Yusef, and Gabriel for our trips in Peru that allowed me to grow and gain life perspective. I thank Nelly who I owe my life to, but departed this life too soon. Lastly, I thank the two women in my life, my wife and my mother, for their unconditional love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

This dissertation consists of three independent essays on international economics. First, I examine how the geographic distribution of innovation determines aggregate productivity. Second, I study how culture matters for production networks and its aggregate implications. Lastly, I focus on how the complementarity of suppliers for production matters for the amplification of negative shocks through production networks.

The first chapter of this dissertation studies the role of spatial knowledge spillovers in R&D for innovation, and therefore on aggregate productivity. I causally estimate spatial knowledge spillovers in Research and Development (R&D) and quantify their importance for R&D policies. Using a new administrative panel on German inventors, I estimate these spillovers by isolating quasi-exogenous variation from the arrival of East German inventors across West Germany after the Reunification of Germany in 1990. I then embed the estimated spillovers into a spatial model of innovation, and use it to quantify the productivity gains from implementing policies that promote R&D activities. The model predicts that reducing migration costs for inventors and R&D subsidies lead to substantial productivity gains. Finally, these productivity gains increase with the degree of spatial knowledge spillovers in R&D.

The second chapter, co-authored with Gaurav Khanna and Hiroshi Toma, examines how cultural proximity shapes production networks, and how it affects welfare. We combine a new dataset of firm-to-firm trade for a large Indian state with information on cultural proximity between firms derived from India's caste and religious classifications. We find that larger cultural proximity between a pair of firms reduces prices and fosters trade at both intensive and extensive margins. We argue that these results are driven by increasing trust between firms due to their cultural proximity, which in turn solves contracting frictions. Guided by these stylized facts, we propose a firm-level production network model, where cultural proximity influences trade and matching costs. Our counterfactual exercises indicate that social inclusion policies raise welfare, and reducing contracting frictions increases welfare via the channel of trade becoming less reliant on cultural proximity.

The third chapter, co-authored with Devaki Ghose and Gaurav Khanna, studies the aggregate implications of firm-level elasticities of substitution across suppliers. We continue using this firm-to-firm trade for a large Indian state, and leverage geographic and temporal variation from the Covid-19 lockdowns in India to estimate these firm-level elasticities of substitution across suppliers of the same product, and quantify the fall in trade. If suppliers 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, suppliers are highly complementary. We 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, and to show the importance of targeted policies during economic downturns according to firm size and connectedness.

Chapter I. Spatial Knowledge Spillovers in R&D and Aggregate Productivity: Evidence from the Reunification of Germany

1.1 Introduction

Research and Development (R&D) is crucial for aggregate productivity due to its direct impact on innovation. At the same time, R&D exhibits substantially higher levels of spatial concentration than overall economic activity. For example, in 2014 in West Germany, a worldwide innovation powerhouse (WEF 2018), around 30% of mechanical engineers worked in the top three cities in this profession. In comparison, only around 18% of workers located in the three most populated cities. Since Marshall (1890), agglomeration economies—spatial and inter-temporal knowledge spillovers, labor pooling, and customer-supplier linkages—are the core explanation for why economic activity concentrates. Nevertheless, the extent, causes and consequences of *spatial knowledge spillovers in R&D*—local productivity gains from the agglomeration of R&D activity—remain elusive. In this paper, I address the following research questions: (i) is there evidence of spatial knowledge spillovers in R&D? and (ii) are they quantitatively important for aggregate productivity when implementing R&D policies?

Addressing these questions is crucial to implement policies that promote economic activity through R&D. Governments around the world implement a variety of policies—reducing mobility or transportation costs, formation of economic clusters, among others—that leverage knowledge spillovers for their effectiveness (Feldman and Kelley 2006). In particular, policies that promote R&D rely strongly on the spatial knowledge spillovers in this sector (Trajtenberg 2001). Moreover, implementing these policies can generate general equilibrium effects due to the internal mobility of agents. Therefore, the design of policies that promote R&D activities requires both well-identified estimates of spatial knowledge spillovers in R&D, and a quantitative framework that accounts for these spillovers in general equilibrium. In this paper, I provide such estimates and framework, and apply them to study policies that promote R&D activities in Germany.

In this paper, I show that spatial knowledge spillovers in R&D are large and important for aggregate productivity. First, using new data on German inventors, I causally estimate such spillovers by isolating quasi-exogenous variation from the arrival of East German inventors across West Germany after the Reunification of Germany in 1990. I find that a 10% increase in the number of inventors in a location leads to average inventor productivity gains of around 4.09%. Second, I build a quantitative spatial model of innovation that account for the spatial knowledge spillovers in R&D I estimated in the data. Third, I calibrate the model and use it to quantify the productivity gains from implementing policies that promote R&D activities. I find that a 25% reduction of migration costs for inventors increases aggregate productivity by 5.87%, and that the 25% subsidy for firms' expenditures in R&D within the 2020 German R&D Tax Allowance Act would increase aggregate productivity by 4.27%. Finally, the productivity gains from these policies increase with the level of spatial knowledge spillovers in R&D. I now describe each of these steps in detail.

In the first part of the paper, I estimate the additional productivity that inventors gain from agglomerating. To perform this task, I leverage a matched administrative data on German inventors between 1980 and 2014. This data exhibits two features that makes it suitable for this paper. First, the dataset includes all the patents and their characteristics that inventors filed over time, so I can calculate the total number of forward citations of inventor's filed patents during a given period—inventor productivity—Second, the dataset tracks how inventors move across locations over time, so I can calculate cluster size as the number of inventors working in a given technological cluster, where a cluster is a technological area-location pair. An example of a cluster is mechanical engineering in Munich.

Then, I leverage variation in cluster size and inventor productivity to estimate the spatial knowledge spillovers in R&D; that is, whether a higher concentration of inventors leads to more productive inventors due to local knowledge spillovers.[1] The analysis compares inventors that moved to clusters of different sizes, and inventors that did not move but the number of inventors in the cluster changed. After saturating the model with a large set of fixed effects, I find that a 10% increase in cluster size is associated with average inventor productivity gains of around 1.75%. This elasticity is statistically significant at the 1%, and its significance is robust to different specifications of inventor productivity and time aggregation.

I then address potential endogeneity concerns that potentially biases when estimating spatial knowledge spillovers in R&D. For example, unobserved inventor idiosyncratic shocks could induce upward or downward biases. For example, if novice inventors systematically sort into large clusters, then spillovers estimates would suffer from downward biases. Also,

[1] Examples of how these spillovers manifest in the real world are interactions and exchange of ideas between inventors (Davis and Dingel 2019).

unobserved cluster shocks could induce upward biases. For example, growth expectations in a technological cluster could increase both the productivity of inventors in that cluster and pull inventors into the cluster, and therefore induce an upward bias. Finally, measurement error could also introduce a downward bias.

To address these concerns, I propose an instrumental variable based on the historical episode of the Reunification of Germany in 1990. In particular, I leverage this natural experiment to construct a shift-share instrument that induces quasi-exogenous variation in the size of West German clusters, which I then use to causally estimate spatial knowledge spillovers in R&D. The "shifts" are leave-out shocks that measure the total number of inventors that moved from each location in East Germany towards any West German cluster, except to the instrumented cluster. The identification assumption is that these shocks are as-good-as-randomly assigned (Borusyak, Hull, and Jaravel 2022); that is, the shifts are uncorrelated with unobservables within the instrumented cluster. The strategy of leaving out the instrumented cluster from the construction of the shifts ensures that the shocks are constructed based solely on push factors arising from the East, and not from pull factors coming from the instrumented cluster. These shocks are then weighted by exposure "shares" that help predicting the number of inventors that move from each East German location to each West German cluster. These shares are constructed based on the inverse of the geographic distance between every location between East and West Germany, and the specialization of each location in East Germany in each technological area. Under this approach, a 10% increase in cluster size leads to average inventor productivity gains of around 4.09%. These spillovers are statistically significant at the 5%, and their significance is robust to different measurements and functional forms for inventor productivity and time aggregation.

In the second part of the paper, to quantify the importance of these spillovers to implement R&D policies, I build a quantitative model of innovation. In each location, a representative firm produces a final good that is consumed locally and is produced by aggregating intermediate inputs from all locations. Each intermediate input is produced by a single firm in each location. Firms hire workers that produce the input, and inventors that engage in R&D. In the model, R&D determines the quality of an input, where firm's inventors generate ideas heterogeneous in productivity, which are then implemented into the firm's blueprint to produce the input at a given quality.

Then, each firm optimally decides how many workers and inventors to hire subject to the demand of its input and its quality. When the firm decides how many inventors to hire, I show that the quality of an input is comprised by the number of ideas a firm's inventors generated, and by how productive these ideas are in expectation. For the first part of input quality, I assume decreasing returns to R&D, so only a subset of firm's inventors generate ideas. This is a valid and necessary assumption since I estimate it in the data, and it is the congestion force that rules out an equilibrium where all inventors move to a single location. For the second part of input quality, following the evidence on spatial knowledge spillovers in R&D and distributional assumptions on the process on how inventors generate ideas, I show that the expected productivity of firms inventors' ideas increases with the spatial knowledge spillovers in R&D in a location.

Finally, I also allow for labor mobility, so workers and inventors choose where to work according to real wages, amenities, and migration costs. And finally, the model allows for straightforward aggregation where aggregate productivity is endogenously determined in general equilibrium. The main prediction of the model is that a location's productivity is endogenously determined by three forces. First, locations with better production fundamentals or that hold more inventors are more productive due to spatial knowledge spillovers in R&D. Second, locations that exhibit higher labor costs are less productive since firms are less able to hire inventors to innovate. Third, locations with higher market access are more productive since higher demand from other locations increases firms' profitability, and therefore their incentive to invest in R&D. All these forces shape location's productivity in general equilibrium. Additionally, the model predicts that a location's productivity acts as an agglomeration force for overall economic activity. Since a location's productivity is determined by the its number of inventors, then locations with more inventors exhibit larger shares in locations' expenditure of intermediate inputs.

In the third part of the paper, I calibrate the model and use it to conduct policy counterfactuals and quantify the importance of spatial knowledge spillovers in R&D for aggregate productivity. I now describe how I discipline the model. First, the model generates an expression that establishes a relationship between inventor productivity and cluster size. This expression is the model counterpart of the specification I used to causally estimate spatial knowledge spillovers in R&D in the data. Then, I can directly import the estimated spillovers into the model. Second, I estimate firm-level decreasing returns to R&D by regressing the number of firm's inventors that filed a patent against the number of hired inventors by the firm. I find an elasticity of 0.65, which confirms the existence of firm-level decreasing returns to R&D. Third, I calibrate migration costs by targeting overall migration rates and estimating migration cost elasticities for both workers and inventors. Finally, I follow Redding (2016) and use aggregate data on wages and the number of workers and inventors across locations to recover unobserved fundamental location productivities and amenities.

After calibrating the model, I conduct counterfactuals to quantify the effect of policies that promote R&D activities on aggregate productivity, and the importance of spatial knowledge spillovers in R&D for the effectiveness of these policies. First, I simulate a supply-side policy of reducing inventor migration costs by 25%. I find that this reduction leads to a 5.87% increase in aggregate productivity. Since the total number of inventors is finite, the policy exhibits substantial heterogeneous effects across locations. I find that the increase in aggregate productivity arises from inventors moving from larger towards smaller clusters in pursue of higher real wages, so the policy reduces the spatial concentration of inventors. Second, I simulate a demand-side policy of a 25% subsidy for firms' R&D expenditure from the 2020 German R&D Tax Allowance Act. I find that this subsidy leads to a 4.27% increase in aggregate productivity. In contrast the reduction of inventor migration costs, all locations increase their productivity and the spatial concentration of inventors increases, so larger clusters exhibit higher productivity gains. Finally, I show that spatial knowledge spillovers in R&D are important for the effectiveness of these policies to foster aggregate productivity.

Literature. This paper contributes to three literature strands. First, this paper contributes to the empirical literature on local knowledge spillovers (Griliches 1991; Jaffe, Trajtenberg, and Henderson 1993; Audretsch and Feldman 1996; Jaffe, Trajtenberg, and Fogarty 2000; Thompson 2006; Carlino, Chatterjee, and Hunt 2007; Combes et al. 2010; Greenstone, Hornbeck, and Moretti 2010; Bloom, Schankerman, and Van Reenen 2013; Kerr and Kominers 2015; Kantor and Whalley 2019; Moretti 2021; Gruber, Johnson, and Moretti 2022). This literature largely focuses on the agglomeration of economic activity, and the positive externalities arising from it. More recently, Moretti (2021) focused in R&D and estimated spatial knowledge spillovers for inventors. I contribute to this literature by exploiting a historical natural experiment to causally estimate spatial knowledge spillovers in R&D.

Second, this paper contributes to the literature on the importance of knowledge spillovers for innovation. This is a vast literature with contributions from urban economics (Eaton and Eckstein 1997; Glaeser 1999; Black and Henderson 1999; Kelly and Hageman 1999; Duranton and Puga 2001; Duranton 2007; Roca and Puga 2017; Duranton and Puga 2019; Davis and Dingel 2019), trade (Ramondo, Rodríguez-Clare, and Saborío-Rodríguez 2016; Hallak and Sivadasan 2013; Atkeson and Burstein 2010; Melitz 2003; Eaton and Kortum 2002; Krugman 1980; Akcigit, Hanley, and Serrano-Velarde 2021), and spatial economics (Desmet and Rossi-Hansberg 2014; Desmet, Nagy, and Rossi-Hansberg 2018; Nagy et al. 2016; Mestieri, Berkes, and Gaetani 2021). I contribute to this literature by building a quantitative framework that explicitly accounts for spatial knowledge spillovers in R&D I estimate in the data.

Third, this paper contributes to the literature on policies that promote productivity and economic growth. This paper focuses on policies that foster labor mobility and R&D activities. This literature show that labor mobility matters for productivity and economic growth both in the data (Borjas and Doran 2012; Burchardi and Hassan 2013; Moser, Voena, and

Waldinger 2014; Peri, Shih, and Sparber 2015; Bosetti, Cattaneo, and Verdolini 2015; Bahar, Choudhury, and Rapoport 2020; Burchardi et al. 2020) and in quantitative settings (Monras 2018; Bryan and Morten 2019; Peters 2022; Arkolakis, Lee, and Peters 2020; Pellegrina and Sotelo 2021; Prato 2021), and that R&D policies can promote productivity (Goolsbee 1998; Romer 2000; Wilson 2009; Acemoglu et al. 2018; Akcigit, Hanley, and Serrano-Velarde 2021). I contribute to this literature by providing a quantitative framework to quantify the productivity gains of implementing migration and R&D policies in general equilibrium.

The remainder of this paper is structured as follows. Section 1.2 explains how I estimate spatial knowledge spillovers in R&D. Section 1.3 describes the model. Section 1.4 maps the model to the data. Section 1.5 presents the results of the counterfactuals. Section 1.6 concludes.

1.2 Spatial Knowledge Spillovers in R&D

In this section I describe the estimation of spatial knowledge spillovers in R&D. The first part of this section describes the data, the second part explains the estimation strategy, and the third part discusses assumptions and results throughout this section. Appendices A.1-A.3 contain additional tables and figures, and details about the data sources.

1.2.1 Data sources

Linked Inventor Biography (INV-BIO). The main dataset in this paper is the INV-BIO by the Research Data Centre of the German Federal Employment Agency at the Institute for Employment Research (FDZ-IAB). The INV-BIO is an administrative dataset comprised by approximately $150,000$ German inventors with high–frequency and detailed information on their employment spells and patenting activities between 1980 and 2014. The INV-BIO is comprised by three modules: (i) an inventor-level module that includes data on inventors' job spells; (ii) an establishment-level module with yearly characteristics of inventors' establishments; and (iii) a patent-level module with information on German inventors' patents. For more details on each of these modules, see Appendix A.3.

Sample of Integrated Employer-Employee Data (SIEED). The FDZ-IAB's SIEED is a 1.5% sample of all establishments in Germany between 1975 and 2018. The dataset tracks establishments' characteristics over time, and establishments' employees' spells over the entire period. I use this complementary dataset to compare the spatial concentration of workers to inventors, and to construct aggregate variables I later use to estimate the model.

1.2.2 Construction of variables.

Dimensions. From the INV-BIO modules I construct an unbalanced panel dataset of inventors. An observation in the data is an inventor i working for establishment ω in location d in technological area a during period t. I focus my analysis on West Germany, which is comprised by 104 labor markets. A labor market is defined based on commuting patterns between districts (Kosfeld and Werner 2012), and are the equivalent to US commuting zones. Finally, to estimate long run estimates of spatial knowledge spillovers in R&D, I stack the data in three 10-year periods: (i) 1982-1991, (ii) 1992-2001, and (iii) 2002-2011.[2]

West German technological clusters. I define a technological cluster as a technological area-location pair. For example, "Mechanical engineering" in "Munich" is a cluster in West Germany. There are 5 technological areas in the data: (i) Electrical engineering, (ii) Instruments, (iii) Chemistry, (iv) Mechanical Engineering, and (v) Others. Then, locations and technological areas comprise $104 \times 5 = 520 \ (d, a)$ technological clusters.

Inventor's cluster. To define an inventor's cluster at a given period, it is necessary to determine the inventor's location and the technological area the inventor works in. First, the location of an inventor is determined by the location of the inventor's establishment since knowledge spillovers arguably happen mostly at the workplace. Additionally, since I consider establishments and not multi-location firms, the location of the inventor is unique. Second, an inventor belongs to the technological area for which he filed the highest share of patents during a given period. For example, if between 1982 and 1991, an inventor filed 80% of his patents in Chemistry, then he belongs to that technological area.

A data limitation is that inventors do not necessarily file patents every period. This generates sample selection, since only inventors that filed a patent during a given period are registered in the data. The main problem arising from this limitation is that it is not straightforward to assign a cluster to an inventor that did not file a patent during a given period. To address this problem, if an inventor did not file a patent during a given period, I assume that an inventor's cluster did not change since since the last time an inventor filed a patent. For example, if in 1995 the latest patent an inventor filed was a Chemistry patent in Dusseldorf in 1993, then I assume that in 1994-1995 the inventor kept working in the Chemistry/Dusseldorf cluster. This is a safe assumption since establishments rarely change locations and inventors tend to specialize in technological areas.

[2] I also consider six 5-year periods to estimate shorter run spatial knowledge spillovers in R&D: (i) 1982-1986, (ii) 1987-1991, (iii) 1992-1996, (iv) 1997-2001, (v) 2002-2006, and (vi) 2007-2011.

Inventor productivity and cluster size. To test for spatial knowledge spillovers in R&D, I construct two main variables. First, I measure inventor productivity $Z^{i\omega}_{da,t}$ as the total number of 5-year forward citations of inventor i's filled patents during period t by the German Patent and Trade Mark Office (DPMA, due to its name in German). If an inventor did not file a patent during period t, then $Z^{i\omega}_{da,t} = 0$. Second, I measure cluster size $R_{da,t}$ as the number of inventors working in cluster (d, a) at the end of period t.

Additional variables. I construct four additional variables I use for both the estimation of spatial knowledge spillovers in R&D in Section 1.2.3 and model calibration in Section 1.4. First, I measure the distance between every location pair $dist_{od}$ as the Euclidean distance (in miles) between the centroids of every labor market in Germany. The district maps were downloaded from the Federal Agency for Cartography and Geodesy, and the correspondence between districts and labor markets is given by Kosfeld and Werner (2012). Second, I measure the technological composition of every location, $TechComp_{da}$, by calculating location d's share of filed patents in technological area a such that $\sum_a TechComp_{da} = 1, \forall d$. Third, I measure migration shares during a given period between every location pair $\left\{\eta^L_{od,t}, \eta^R_{od,t}\right\}$ for workers and inventors, respectively. Fourth, I measure average wages in a given period for every location $\left\{w^L_{o,t}, w^R_{o,t}\right\}$ for workers and inventors, respectively.

1.2.3 Estimation

OLS estimates To measure spatial knowledge spillovers, I consider the following specification between inventor productivity $Z^{i\omega}_{da,t}$ and cluster size $R_{da,t}$:

$$\log\left(Z^{i\omega}_{da,t}\right) = \iota_{d,t} + \iota_{a,t} + \iota_{da} + \iota_{\omega} + \iota_i + \beta \log\left(R_{da,t}\right) + \epsilon^{i\omega}_{da,t}. \tag{1.1}$$

If there are spatial knowledge spillovers in R&D, then $\beta > 0$. I saturate the model with a large set of fixed effects. $\iota_{d,t}$ are location/period fixed effects that account for amenities and location shocks that drive the overall activity of a location. $\iota_{a,t}$ are technological area/period fixed effects that account for overall technological shocks. ι_{da} are cluster fixed effects that account for time-invariant cluster productivity, and for the fact that some clusters file more patents than others in average. ι_{ω} are establishment fixed effects that account for inventor sorting due to time-invariante establishment productivity. ι_i are inventor fixed effects that control for inventor sorting due to time-invariant inventor productivity. In all specifications, standard errors are clustered at the (d, a) level. The identification assumption is that inventor unobservables $\epsilon^{i\omega}_{da,t}$ are uncorrelated with cluster size $R_{da,t}$.

The main measurement challenge is to account for zeros in the dependent variable $Z^{i\omega}_{da,t}$. I consider $\log\left(1 + Z^{i\omega}_{da,t}\right)$ as the dependent variable for the main specifications. Table 1 report the OLS estimates of Equation (1.1). Columns $(1) - (6)$ show the value of the estimated spillovers as I progressively include the aforementioned fixed effects. The value of these estimates remain around 0.12. Column (6) reports the main OLS estimate that includes inventors fixed effects, which is key to compare a given inventor across periods and clusters. This estimate indicates that an inventor whose cluster size increased by 10% or moved to a cluster with 10% more inventors reports productivity gains of 1.75% in average.

Table 1: OLS models

	(1)	(2)	(3)	(4)	(5)	(6)
$\log\left(R_{da,t}\right)$	0.0705	0.111	0.0985	0.109	0.0896	0.175
	(0.0256)	(0.0170)	(0.0166)	(0.0385)	(0.0358)	(0.0660)
$\iota_{d,t}$		✓	✓	✓	✓	✓
$\iota_{a,t}$			✓	✓	✓	✓
ι_{da}				✓	✓	✓
ι_{ω}					✓	✓
ι_{i}						✓
N	$177,301$	$177,300$	$177,300$	$177,294$	$162,803$	$84,639$
R^2	0.008	0.053	0.064	0.079	0.246	0.700

Notes: In this Table I report OLS estimates from Equation (1.1). The dependent variable is measured as $\log\left(1 + Z^{i\omega}_{da,t}\right)$, and $Z^{i\omega}_{da,t}$ is the number of 5-year forward citations from the DPMA. The table is comprised by 6 columns. Each column corresponds to a different combination of fixed effects, as pointed out by rows $4-8$. Row 2 reports the estimate of β. Row 3 reports standard errors clustered at the (d, a) level. Rows $9 - 10$ report the number of observations and the goodness of fit, respectively.

Robustness. Table A2 contains the estimated spillovers under different specifications of inventor productivity. Since column (6) is the main specification in Table 1 I focus the robustness discussion around this specification. Panel A shows results when patent citations arose from the European Patent Office (EPO) and the EU (both the DPMA and EPO), respectively. Under these specifications, column (6) shows that the elasticities of inventor productivity to cluster size are 0.173 and 0.245, respectively. These spillovers are comparable to the ones reported in Table 1. Panel B shows results when I account for zeros by using the Inverse Hyperbolic Sine (IHS) for inventor citations instead of $\log(1 + Z)$. Column (6) shows that the elasticity of inventor productivity to cluster size is 0.217, which is similar to the baseline estimate of 0.175. Additionally, when patent citations arose from the EPO and the EU, elasticities are around $0.21 - 0.24$.

Finally, results also hold under shorter time horizons. In Table A3, I show the estimated spillovers when the frequency of the data is 5-year periods, where the first row measures inventor productivity as $\log(1+Z)$, and the third row measures it as $IHS(x)$. In both cases, column (6) shows that the spillovers are around 0.1, so the magnitude of spatial knowledge spillovers in R&D scale with the frequency of the data. The intuition for these results is that longer time horizons allow for larger spillovers to manifest in the data.

1.2.4 IV approach

Endogeneity concerns. To causally estimate β from Equation (1.1), the key identification assumption is that the unobservables $\epsilon^{i\omega}_{da,t}$ are uncorrelated to cluster size $R_{da,t}$. Nevertheless, there are at least two endogeneity concerns that could potentially violate this assumption. First, unobserved time-varying idiosyncratic shocks can bias the estimate of β. For example, inventors can decide to start working in a given technological area due to unobservable reasons. If inventors at the beginning of their careers in a given technology, who initially report low productivity, move to large clusters due to better career prospects, this would introduce a downward bias on β. On the other side, if inventors at the peak of their careers in a given technology, who report high productivity, move to large clusters due to even better career prospects, this would introduce an upward bias on β.

Second, unobserved time-varying cluster-level shocks can introduce an upward bias when estimating β. For example, a sudden increase in growth expectations for Chemistry in Dusseldorf could increase both cluster size due to an inflow of inventors towards that cluster, and inventor productivity in that cluster, introducing an upward bias on β. Finally, measurement error could also bias the estimate of β downwards. To address these endogeneity concerns, I then propose an instrumental variable approach to causally estimate β. In summary, I leverage quasi-exogenous variation in cluster size arising from the arrival of East German inventors towards West German clusters during the Reunification of Germany in 1990.

Brief historical background: The Reunification of Germany. During the final phase of World War II, the Potsdam Agreement was signed between the US, the UK, and the USSR on August 1st 1945. Part of this agreement was the division of Germany in two main blocs: (i) the Federal Republic of Germany (FRG, also known as "West Germany"), and (ii) the German Democratic Republic (GDR, also known as "East Germany"). FRG was based on liberal economic-social institutions from the West, while GDR was based on socialist institutions from the ex-Soviet Union.

In 1952, the borders between East and West Germany were well-established. Nevertheless, migration was still allowed between the two blocs. This lasted until 1961, when migration between these two blocs ceased. Then, in October 3rd 1990, the GDR was dissolved and the process to reunify Germany began. During this period, the "Exodus to the

West" started, where a large number of East Germans migrated to the West. Figure A2 plots the magnitude of this shock, which was considered to be unexpected and be permanent at the time. Since inventors from East Germany also moved to the West (Hoisl et al. 2016), I use the variation arising from the arrival of East German inventors across West German clusters.

IV estimates. To motivate the design of my instrument, consider an ideal experiment to causally estimate spatial knowledge spillovers in R&D. In this thought experiment, I would randomize inventors' clusters in West Germany, such that productivity gains arising from changes in cluster size can be estimated. Since it is not possible to obtain such exogenous variation, I extract quasi-exogenous variation in cluster size from the Reunification of Germany. To do this, I construct a shift-share instrument based on the arrival of East German inventors across West German clusters. If the variation in cluster size arising from the overall arrival of East German inventors is as-good-as-random, then this is sufficient to causally estimate spatial knowledge spillovers in R&D. First, I use variation in the arrival of inventors towards West German clusters, so the second stage regression in first differences of Equation (1.1) is

$$\Delta \log\left(Z^{i\omega}_{da,t}\right) = \iota_{d,t} + \iota_{a,t} + \beta \Delta \log\left(R_{da,t}\right) + \Delta \epsilon^{i\omega}_{da,t}. \tag{1.2}$$

Notice that the fixed effects in Equation (1.2) that prevail after introducing first-differences are location/period $\iota_{d,t}$ and technological area/period $\iota_{a,t}$ fixed effects. $\iota_{d,t}$ are crucial to control for the overall arrival of East Germans to West German locations during the Reunification. Also, $\iota_{a,t}$ accounts for overall technological change that could have happened during Reunification. Now, the first stage regression is a shift-share instrument:

$$IV_{da,t} = \sum_{o \in \mathcal{E}} g_{o,t} \times s_{o,da}, \tag{1.3}$$

where $o \in \mathcal{E}$ is location o in East Germany ($\mathcal{E}$), and d is a location in West Germany. The instrument is constructed as the interaction of two terms: (i) a common set of shocks to West German clusters $g_{o,t}$ (i.e. the "shifts"); and (ii) a set of exposure weights to these shocks $s_{o,da}$ (i.e. the "shares"). The shifts $g_{o,t} \equiv \log\left(\Delta R^{-d,-a}_{o,t}\right)$ are the log of the number of inventors in o that moved to any West German cluster except the instrumented cluster (d, a) during period t. Following Borusyak, Hull, and Jaravel (2022), the identification assumption to estimate β is that the overall arrival of East German inventors in West Germany $g_{o,t}$ excluding the instrumented cluster is as-good-as-random. That is, the shifts are uncorrelated with inventor unobservables within the instrumented cluster. This is a safe assumption since the instrumented cluster is being left out to construct each shift, so the shifts are constructed based solely on push factors arising from each East German location, and are clean from pull factors coming from the instrumented cluster.

The shifts are then weighted by exposure shares, which help predicting how many inventors from each East German cluster will move to each West German cluster. The shares $s_{o,da} \equiv dist_{o,d}^{-1} \times TechComp_{o,a}$ are comprised by two terms: (i) $dist_{o,d}^{-1}$ is the inverse distance between o and d; and (ii) $TechComp_{o,a}$ is the technological composition of location o. The construction of these variables is detailed in Section 1.2.2. The intuition of the shares is the following. First, migration flows decay with distance, so locations closer to each other should exhibit higher migration shares. This is consistent with Hoisl et al. (2016) who find that distance was indeed a key predictor for the migration from the East to the West. Second, the specialization of East German locations towards different technologies predicting which technological area an East German inventor will work on upon moving to the West. The shares are then normalized such that $\sum_{o \in \mathcal{E}} s_{o,da} = 1, \forall d, a$.

Table 2 contains the IV estimates of spatial knowledge spillovers in R&D. All the estimates exhibit an F-statistic above 10, which reflects the relevance of the proposed instrument. Column (1) reports the estimate of the spillovers when I do not consider any fixed effects. This reports a value of 0.178 which is similar to the OLS estimate from column (6) in Table (1). It is crucial to include location-period fixed effects to account for the overall arrival of East Germans to West Germany. In column (2) I show that estimated spillovers after including these fixed effects are 0.309. Finally, it is also key to include technological area-period fixed effects to control for technological changes after the Reunification. Column (3) contains the main empirical result of this paper: an inventor whose cluster size increases by 10% or moved to a cluster with 10% more inventors becomes 4.09% more productive in average. This estimate is between 2 and 3 times the OLS estimate of 1.75% from column (6) in Table (1), which reflects a downward bias when estimating β due to unobservables and measurement error.

Table 2: IV models

	(1)	(2)	(3)
$\Delta \log (R_{da,t})$	0.178	0.309	0.409
	(0.0431)	(0.101)	(0.152)
$\iota_{d,t}$		✓	✓
$\iota_{a,t}$			✓
$KP-F$	132.1	34.14	28.23
N	$50,778$	$50,776$	$50,776$

Notes: In this Table I report IV estimates from Equation (1.2), where the instrument is constructed as in Equation (1.3). The dependent variable is measured as $\Delta \log \left(1+Z_{da,t}^{i\omega}\right)$, and $Z_{da,t}^{i\omega}$ is the number of 5-year forward citations from the DPMA. The table is comprised by 4 columns. Each column corresponds to a different combination of fixed effects, as pointed out by rows $5-6$. The fourth column reports the OLS estimate from Equation (1.2). Row 3 reports the estimate of β. Row 4 reports standard errors clustered at the (d,a) level. Rows $7-8$ report the first stage Kleibergen-Paap F-statistic (KP-F) and the number of observations, respectively.

Robustness. Table A4 contains the estimated spillovers under different specifications of inventor productivity. Since column (3) is the main specification in Table 2, I focus the robustness discussion around this specification. Panel A shows results when patent citations arose from the EPO and the EU, respectively. Under these specifications, column (3) shows that the elasticities of inventor productivity to cluster size are 0.209 and 0.343, respectively. These spillovers are comparable but somewhat lower to the ones reported in Table 2. Panel B shows results when I account for zeros by using the IHS for inventor citations instead of $\log(1+Z)$. Column (3) shows that the elasticity of inventor productivity to cluster size is 0.498. Additionally, when patent citations arose from the EPO and the EU, elasticities are around $0.23-0.39$.

Finally, results also hold under shorter time horizons. In Table A5 I show the estimated spillovers when the frequency of the data is 5-year periods, where the first row measures inventor productivity as $\Delta \log(1+Z)$, and the third row measures it as $\Delta IHS(x)$. In both cases, column (3) shows that the spillovers are around 0.09, so the magnitude of spatial knowledge spillovers in R&D scale with the frequency of the data. The intuition for these results is that longer time horizons allow for larger spillovers to manifest in the data.

1.2.5 Discussions

Do citations measure productivity? Throughout this paper, I have measured inventor productivity as the number of forward citations of all inventor's filed patents during a given period. Then, it is reasonable to pose whether number of citations indeed measure productivity. There is a vast literature that documents a positive relationship between number of citations and proxies for productivity, such as patent value (Kogan et al. 2017; Hall, Jaffe, and Trajtenberg 2001; Harhoff et al. 1999; Trajtenberg 1990).

More recently, Abrams, Akcigit, and Grennan (2013) find preliminary evidence of a inverse U-shaped relationship between number of citations and patent value in the data. They rationalize this finding by distinguishing between productive and strategic patents. For the former, more citations reflect a higher patent productivity since a citation reflects further creation of patents. For the latter, patenting an idea maintain incumbent's monopoly power such that entry is inhibited, so the number of citations decreases. To check whether German citations are mostly productive or strategic, I review literature on firm surveys about their incentives to patent (Blind et al. 2006; Cohen et al. 2002; Pitkethly 2001; Duguet and Kabla 2000; Schalk, Tager, and Brander 1999; Arundel, Paal, and Soete 1995), which is mostly focused on Europe, particularly Germany. In general, the major motive for German firms to file patents is the classical incentive to protect their ideas, which goes in line with productive patenting.

Is it exposure instead of knowledge spillovers? A possible identification threat to estimate β is that the number of citations reflect higher exposure of an inventor's ideas, which is orthogonal to knowledge spillovers. For example, if an inventor moves to a larger cluster, then his ideas could obtain more exposure to a larger share of inventors, so his patents get cited more often. This would introduce an upward bias when estimating β. I present two main arguments against this concern.

First, the patenting market is drastically different from other industries that rely on citations, such as academia. In academia, citations measure aspects other than productivity such as reputation, exposure, among others. In the patenting market, citations are required whenever an invention uses information from another patent. Whenever a citation this situation does not take place, a patent infringement has taken place, so then the owner of the non-cited patent can pursue legal means to resolve the issue. This is particularly relevant for the industrial economy of Germany that reports one of the largest number of patent litigation cases (Cremers et al. 2017), and exhibits one of the highest cross-country levels of patent enforcement (Papageorgiadis and Sofka 2020).

Second, assuming that these effects are biasing the estimate of β, Tables A2 and A4 include the OLS and IV estimates where productivity is measured by the number of citations from the EPO, which is the European patenting institution and completely independent from the German patenting office. These estimates still provide evidence on the existence of spatial knowledge spillovers in R&D.

Comparison to previous estimates. I now compare my estimates with previous literature. Carlino, Chatterjee, and Hunt (2007) shows that the rate of patenting per capita is around 1.95% higher in a US metropolitan area with 10% higher population density. My baseline estimate of 4.09% is higher due to three differences. First, I test for knowledge spillovers in R&D by measuring productivity through number of citations instead of patenting rates. Second, I estimate long-run spatial knowledge spillovers since I consider 10-year periods. In contrast, they leverage cross-sectional variation across US metropolitan areas. Third, my identification relies on a historical natural experiment instead on the inclusion of covariates.

Moretti (2021) is the closest to this paper. His OLS estimate is around 0.067, while my estimate from Table 1 is 0.175. When running the model in first differences, his IV estimate is around 0.049, while my estimates from Table 2 is 0.409. Even thought both of these papers estimate spatial knowledge spillovers in R&D at the inventor level, the differences in magnitudes arise due to two differences. First, I estimate long-run spillovers (10-year periods), while Moretti estimates short-run spillovers (1-year periods). Second, my larger estimates could result from stronger spatial knowledge spillovers in R&D in Germany in comparison to the US.

1.3 Model

In this section I build a quantitative spatial model of innovation. Appendices A.4-A.5 contain details about the derivations in the model.

1.3.1 Setup

Geography. There is a discrete set of locations $\mathcal{S} \equiv \{1, 2, \ldots, S\}$, where $o \in \mathcal{S}$ is the origin location, and $d \in \mathcal{S}$ is the destination location. When I take the model to the data, I consider $\mathcal{S}$ to be the 104 labor markets in West Germany.

Firms. There are two types of firms in each location: (i) a final good firm, and (ii) a unit mass of intermediate input firms. The final good is produced by a representative firm, it is non-tradable, and it is produced by aggregating intermediate inputs from all locations with constant elasticity of substitution (CES). Each input is produced by a single firm, it is tradable across locations, it is produced by firm's workers, and its quality is determined through R&D by firm's inventors. Because the mass of firms is fixed in each location, firms earn positive profits, which in turn are invested in a national fund and redistributed proportionally to all agents in the economy.

Agents. There are two types of agents: (i) workers, and (ii) inventors. Each agent supplies a unit of labor inelastically, earns income from their wage and redistributed profits, and consumes the local final good. Finally, both workers and inventors are mobile, so they optimally decide where to work by maximizing their utility subject to migration costs.

1.3.2 Technology

Final good firms. In each location d, a representative firm produces a final good by aggregating intermediates from all locations. The production function of the final good is

$$Q_d = \left(\sum_o \int_{\omega \in \Omega_{od}} A_o^{\omega \frac{1}{\sigma}} Q_{od}^{\omega \frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \tag{1.4}$$

where Ω_{od} is the set of intermediate input firms in o selling to d, Q_d is the production of the final good, Q_{od}^{ω} is the quantity of intermediate input ω from o sold to the final good firm in d, A_o^{ω} is the quality of input ω from o, and $\sigma > 1$ is the CES across intermediate inputs. The final good firm maximizes profits subject to Equation (1.4), which yields the demand for intermediate inputs

$$Q_{od}^{\omega} = A_o^{\omega} P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d, \tag{1.5}$$

where $P_d^{1-\sigma} = \sum_o P_{od}^{1-\sigma}$ and $P_{od}^{1-\sigma} = \int_{\omega \in \Omega_{od}} A_o^{\omega} P_{od}^{\omega^{1-\sigma}} d\omega$ are CES price indices, and $X_d = P_d Q_d$ is total expenditure on the final good in d.

Intermediate input firms. In each location o, there is a unit mass of firms, where each produce a unique and tradable intermediate input. The profits of firm ω selling to d is

$$\pi_{od}^{\omega} = P_{od}^{\omega} Q_{od}^{\omega} - \tau_{od} w_o^L L_{od}^{\omega}, \tag{1.6}$$

where w_o^L are worker wages, L_{od}^{ω} is labor demand by ω, and $\tau_{od} > 1$ are iceberg trade costs. A unit of labor is required to produce an intermediate input:

$$L_{od}^{\omega} = Q_{od}^{\omega}. \tag{1.7}$$

Then, firm ω maximizes total profits subject to Equations (1.5), (1.6), and (1.7):

$$\begin{aligned} \max_{\left\{P^{\omega}_{od}, Q^{\omega}_{od}, L^{\omega}_{od}\right\}} \quad \pi^{\omega}_{o} &= \sum_{d} \pi^{\omega}_{od} \\ s.t.& \\ \pi^{\omega}_{od} &= P^{\omega}_{od} Q^{\omega}_{od} - \tau_{od} w^{L}_{o} L^{\omega}_{od}, \\ L^{\omega}_{od} &= Q^{\omega}_{od}, \\ Q^{\omega}_{od} &= A^{\omega}_{o} P^{\omega\,-\sigma}_{od} P^{\sigma-1}_{d} X_{d}. \end{aligned} \tag{1.8}$$

Then, firms charge a constant markup:

$$P^{\omega}_{od} = \overline{m} \tau_{od} w^{L}_{o}, \forall \omega \in \Omega_{od} \tag{1.9}$$

where $\overline{m} \equiv \frac{\sigma}{\sigma-1}$ is the CES constant markup over marginal costs. Plugging back Equation (1.9) in (1.8), firm total profits are

$$\pi^{\omega}_{o} = \frac{1}{\sigma} A^{\omega}_{o} \sum_{d} \left(\frac{P^{\omega}_{od}}{P_{d}}\right)^{1-\sigma} X_{d}. \tag{1.10}$$

From (1.10), we notice that total profits of firm ω increase with market demand from every location d and the quality of its intermediate input A^{ω}_{o}. This is because inputs of higher quality exhibit higher demand from every final good firm.

Quality of intermediate inputs. Each firm ω in every location owns a blueprint that describes the production process of intermediate input with quality A^{ω}_{o}. The blueprint is comprised by n^{ω}_{o} ideas generated by firm's inventors, and ideas are heterogeneous in productivity. Then, the quality of the intermediate input is

$$A^{\omega}_{o} = \mathbb{Z}^{\omega}_{o} n^{\omega}_{o}, \tag{1.11}$$

where $\mathbb{Z}^{\omega}_{o}$ is the expected productivity of inventors' ideas. In Appendix A.4.2 I provide two microfoundations that generate isomorphic expressions for the quality of intermediates up to a constant. To provide intuition on how the quality of the intermediate input is determined by the expected productivity of inventors' ideas, I briefly sketch the first microfoundation based on necessary tasks. Consider that a firm ω owns a blueprint that contains a continuum of tasks to produce a unit of its input. The firm hires R^{ω}_{o} inventors who produce $n^{\omega}_{o} \leq R^{\omega}_{o}$ ideas to be implemented in its blueprint, and these ideas are heterogeneous in productivity. Given the assumption that each of these ideas improve the quality of every task within the firm's blueprint, and all tasks are necessary to produce a unit of the firm's input, then the expected productivity of the implemented ideas into the blueprint captures the overall quality of the firm's input. I also consider decreasing returns to R&D, such that

$$n_o^{\omega} = R_o^{\omega^{\zeta}}, \tag{1.12}$$

where $\zeta \in (0,1)$ is the degree of decreasing returns to R&D. This assumption introduces a local congestion in R&D, which is key to countervail agglomeration forces in R&D, and therefore avoid the possibility of all inventors concentrating in a single location. In Section 1.4 I estimate ζ and provide empirical evidence that there are indeed firm-level decreasing returns to R&D in the data. Now, what is left to is model the process for which firm's inventors generate ideas, which in turn determine $\mathbb{Z}_o^{\omega}$.

Productivity of ideas. An inventor i hired by firm ω generates an idea to be implemented into the firm's blueprint on how to produce a unit of the firm's input. Ideas are heterogeneous in productivity $Z_o^{i\omega}$ drawn from a probability distribution:

$$Z_o^{i\omega} \sim Frechet\left(\alpha, \lambda_o^{\frac{1}{\alpha}}\right), \tag{1.13}$$

where α and $\lambda_o^{\frac{1}{\alpha}}$ are the shape and scale parameters of the Frechet distribution, respectively. Appendix A.4.1 describes inventors' innovation process based on Kortum (1997) that generates a Frechet distribution for the productivity of inventors' ideas. Under this framework, λ_o is referred as the *spillover function* since it embeds exogenous economic forces that increase inventors' productivity. Guided by the empirical evidence on spatial knowledge spillovers in R&D in Section 1.2.3, I consider the following functional form:

$$\lambda_o^{\frac{1}{\alpha}} = \mathcal{A}_o R_o^{\tilde{\gamma}}, \tag{1.14}$$

where $\mathcal{A}_o$ is a fundamental location productivity, R_o is the number of inventors in o (i.e. *cluster size*), and $\tilde{\gamma} \equiv \frac{\gamma}{\alpha}$ are spatial knowledge spillovers in R&D.[3] Finally, considering the probability distribution in Equation (1.13), then the expected productivity of inventors' ideas is

$$\mathbb{Z}_o^{\omega} = \psi \lambda_o^{\frac{1}{\alpha}}, \tag{1.15}$$

where $\psi > 0$ is a constant that arises from the microfoundation for the quality of intermediate inputs.

[3]Technically, γ are spatial knowledge spillovers in R&D. Since γ and α are not separable, I consider $\tilde{\gamma} \equiv \frac{\gamma}{\alpha}$ to denote spatial knowledge spillovers in R&D throughout the paper.

Research and Development (R&D). Firms ω in every location engages in R&D, who optimally decide how many inventors to hire. The optimal number of inventors arises from the trade-off between the cost of hiring inventors and higher quality. Then, firm ω maximizes total profits after R&D expenditure subject to (1.10), (1.14), (1.11), and (1.15):

$$\max_{\{R_o^\omega\}} \overline{\pi}_o^\omega = \pi_o^\omega - w_o^R R_o^\omega \tag{1.16}$$

$$s.t.$$

$$\pi_o^\omega = \frac{1}{\sigma} A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d,$$

$$A_o^\omega = \psi \mathcal{A}_o R_o^{\widetilde{\gamma}} R_o^{\omega\zeta}.$$

Then, firms' demand for inventors is

$$R_o^\omega = \left(\frac{\zeta}{\sigma}\frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}. \tag{1.17}$$

1.3.3 Location choice

In each location d, there are two types of agents: inventors ($n = R$), and workers ($n = L$). Upon moving to d, agents maximize their utility subject to their budget constraint. Agents have preferences for consuming local final goods and location amenities. Then, the agents' indirect utility is

$$U_d^n = \frac{\mathcal{B}_d^n w_d^n \left(1 + \overline{\pi}\right)}{P_d} \quad, n = \{L, R\}, \tag{1.18}$$

where $\mathcal{B}_d^n$ are type-specific location amenities, and $\overline{\pi}$ are redistributed profits per-capita. Then, an agent i of type n working in o moves to d by maximizing its utility:

$$U_{od}^{i,n} = \max_{d \in \mathcal{S}} \left\{\frac{U_d^n}{\mu_{od}^n} \times \epsilon^i\right\} \quad, n = \{L, R\}, \tag{1.19}$$

where $\mu_{od}^n > 1$ are type-specific *iceberg* migration costs, $G(\epsilon) = \exp\left(-\epsilon^{-\kappa}\right)$ are location preference shocks, and κ is the spatial labor supply elasticity. Then, using the order-statistic properties of the Frechet distribution from Equation (1.19), the share of agents of type n moving from o to d is

$$\eta_{od}^n = \frac{\left(\frac{U_d^n}{\mu_{od}^n}\right)^\kappa}{\sum_\delta \left(\frac{U_\delta^n}{\mu_{o\delta}^L}\right)^\kappa} \quad, n = \{L, R\}. \tag{1.20}$$

1.3.4 Aggregate variables

Aggregate productivity. In the model, a location's productivity is measured as the average quality of intermediates in a location. Due to firm symmetry and the fixed mass of firms in each location, from Equations (1.14)-(1.12) and (1.17), location's productivity is

$$A_o^{1-\zeta} \propto \underbrace{\left(\mathcal{A}_o R_o^{\widetilde{\gamma}}\right)}_{spillovers} \underbrace{\left(w_o^{L^{\sigma-1}} w_o^R\right)^{-\zeta}}_{labor\ costs} \underbrace{\left(\sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d\right)^{\zeta}}_{market\ access}. \tag{1.21}$$

Equation (1.21) is the main expression of the model since it encapsulates the market forces that determine aggregate productivity in general equilibrium, and what different policies imply for productivity. First, locations that exhibit higher local spillovers are more productive. On one side, there are some locations that are fundamentally more prone for innovation than others ($\mathcal{A}_o$). More interestingly, due to spatial knowledge spillovers in R&D, locations that exhibit higher R&D employment (R_o) are more productive. Second, locations that report higher labor costs $\left\{w_o^L, w_o^R\right\}$ are less productive because higher labor demand increases overall labor costs, which reduces firm hiring, and therefore pushes R&D incentives down.

These two forces reflect the main trade-off when implementing different R&D policies. Demand-side R&D policies, such as R&D subsidies, can increase a location's productivity since it increases the number of inventors in a location, but i also increases labor costs. In the end, the net effect on productivity depends on which force prevails in equilibrium (Gruber, Johnson, and Moretti 2022). On the other side, supply-side policies, such as the reducing migration costs for inventors, can drastically increase the productivity of a location since it increases number of inventors and reduce hiring costs. By the same argument, these policies can also have drastic distributional effects due to inventors reallocation.

Finally, locations with higher market access are more productive. This term arises due to the tradability of intermediate inputs. Intuitively, higher demand for a firm's inputs pushes its to engage in R&D and increase the quality of its input.

Price indices. Given Equations (1.9) and (1.21), price indices are

$$P_{od}^{1-\sigma} = A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma} \quad and \quad P_d^{1-\sigma} = \sum_o A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}. \tag{1.22}$$

Trade shares. Given Equations (1.22), location o's share in location d's expenditure is

$$\chi_{od} = \frac{A_o \left(\tau_{od} w_o^L\right)^{1-\sigma}}{\sum_o A_o \left(\tau_{od} w_o^L\right)^{1-\sigma}}. \tag{1.23}$$

Equation (1.23) shows that trade shares depend directly on locations' productivity A_o; that is, a higher A_o increases location o's share in d's total expenditure. Intuitively, the higher quality of intermediate inputs from o increases their demand from all other locations, which in turn increases o's trade share. Moreover, some of the concentration of economic activity is explained by agglomeration forces arising from locations' R&D activity.

Profits per-capita. Firms' profits are invested in a national fund, and they are then redistributed uniformly across the country's population. Then, plugging Equation (1.17) back in firm profits (1.16), yields location's total profits:

$$\overline{\pi}_o = \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma} \psi}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R\zeta} w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}. \tag{1.24}$$

Then, profits per-capita are

$$\overline{\pi} = \frac{1}{N} \sum_o \overline{\pi}_o. \tag{1.25}$$

1.3.5 Equilibrium

Inventors market. From firm's demand for inventors (1.17), location o's aggregate demand for inventors is

$$w_o^R = \frac{\zeta \psi \overline{m}^{1-\sigma}}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}-(1-\zeta)}}{w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d. \tag{1.26}$$

Notice in Equation (1.26) that, if $\widetilde{\gamma}$ is sufficiently high, then the demand for inventors is upward sloping. That is, if spatial knowledge spillovers in R&D are higher than the decreasing returns to R&D, then the demand for inventors will exhibit an upward slope. This is a similar mechanism as in Allen and Donaldson (2020) and Krugman (1979), where a sufficiently strong productivity spillovers can lead to non-unique equilibria. I later show that the calibration of the model rules out the possibility of multiple equilibria. Now, given Equation (1.20), the supply of inventors in each location is

$$R_d = \sum_o \eta_{od}^R \overline{R}_o, \tag{1.27}$$

where $\{\overline{R}_o\}_{\forall o \in \mathcal{S}}$ is the exogenous allocation of inventors across locations.

Workers market. From Equations (1.5), (1.7), and (1.9), the aggregate demand for workers is

$$w_o^L = \overline{m}^{-1} \left(\frac{A_o}{L_o} \sum_d \tau_{od}^{-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{\sigma}}. \tag{1.28}$$

Given Equation (1.20), the equilibrium number of workers and inventors in each location is

$$L_d = \sum_o \eta_{od}^L \overline{L}_o, \tag{1.29}$$

where $\{\overline{L}_o\}_{\forall o \in \mathcal{S}}$ is the exogenous allocation of workers across locations.

Equilibrium in goods market. To close the model, in every location, total income equals total expenditure. Income Y_o is comprised by wages earned by workers and inventors:

$$Y_o = (1 + \overline{\pi}) \left(w_o^L L_o + w_o^R R_o \right). \tag{1.30}$$

Expenditure X_o is comprised by purchased intermediates from every location d:

$$X_o = \sum_d \chi_{od} X_d. \tag{1.31}$$

In equilibrium, given Equation (1.23), income equals expenditure $X_o = Y_o, \forall o$:

$$w_o^L L_o + w_o^R R_o = \sum_d \chi_{od} \left(w_d^L L_d + w_d^R R_d \right) \tag{1.32}$$

Government budget. A national government implements a set of subsidies s_o that are funded with a uniform labor tax τ. I assume a balanced government budget constant, such that

$$\tau \sum_o \left(w_o^L L_o + w_o^R R_o \right) = \sum_o s_o \left(w_o^R R_o \right). \tag{1.33}$$

Definition 1 (Equilibrium). *Given iceberg trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, iceberg migration costs $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, location fundamentals $\{\mathcal{A}_o, \mathcal{B}_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$, an equilibrium is a set of wages $\{w_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$, prices $\{P_o\}_{\forall o \in \mathcal{S}}$, quantities $\{L_o, R_o, Q_o\}_{\forall o \in \mathcal{S}}$, and aggregate productivity $\{A_o\}_{\forall o \in \mathcal{S}}$ such that (i) inventor markets clear (Equations (1.26) and (1.27)), (ii) worker markets clear (Equations (1.28) and (1.29)), (iii) goods market clear (Equation (1.32)), and (iv) locations' productivity are determined by (1.21).*

1.3.6 Equilibrium with R&D subsidies

To evaluate the implementation of R&D subsidies, I consider a set of subsidies s_o that are funded with a uniform labor tax τ. First, firms' total profits after R&D is $\overline{\pi}_o^\omega = \pi_o^\omega - (1 - s_o)\, w_o^R R_o^\omega$, so the subsidy acts as a positive shock for the demand of inventors which value. Second, since these subsidies are funded through a uniform labor tax, then workers and inventors' income are now $w_o^n (1 + \overline{\pi} - \tau)$. Finally, I assume a balanced government budget constant, so government's income $\tau \sum_o \left(w_o^L L_o + w_o^R R_o \right)$ equals its expenditure $\sum_o s_o \left(w_o^R R_o \right)$.

1.4 Taking the Model to the Data

In this section I describe the calibration strategy of the model. The model is parametrized by spatial knowledge spillovers in R&D $\{\widetilde{\gamma}\}$, decreasing returns to R&D $\{\zeta\}$, type-specific migration costs $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, fundamental location productivity $\{\mathcal{A}_o\}_{\forall o \in \mathcal{S}}$, type-specific fundamental location amenities$\{\mathcal{B}_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, and remaining parameters $\{\alpha, \kappa, \sigma\}$. Table 5 at the end of this section summarizes the calibration strategy of the model, and further details on the parametrization are in Appendix A.6.

Spatial knowledge spillovers in R&D $\{\widetilde{\gamma}\}$. The reduced-form estimates for spatial knowledge spillovers in R&D in Section 1.2 are mapped to $\widetilde{\gamma}$. Consider Equation (1.15), which describes how cluster size increases the expected productivity of inventors. Considering Equation (1.14), the model yields a log-log relationship between inventor productivity and cluster size:

$$\log \left(Z_o^{i\omega} \right) = \iota + \iota_o + \widetilde{\gamma} \log \left(R_o \right) + \epsilon_o^{i\omega}, \tag{1.34}$$

where $\iota \equiv \log(\psi)$ and $\iota_o \equiv \log(\mathcal{A}_o)$. After considering the additional time dimension t and technological areas a, and first differences, Equation (1.34) is the model counterpart of Equation (1.2) which was used to estimate spatial knowledge spillovers in R&D $\beta = 0.409$. Notice that, technically, β is the elasticity of inventor 5-year forward citations to cluster size, while $\widetilde{\gamma}$ is the elasticity of patent/idea productivity or quality to cluster size. Therefore, the value of $\widetilde{\gamma}$ is such that $\widetilde{\gamma} = \delta\beta$, where δ is the elasticity of patent/idea productivity or quality to 5-year forward citations. I follow Lanjouw and Schankerman (2004) and consider $\delta = 0.22$, such that$\widetilde{\gamma} = \delta\beta = (0.22)(0.409) \approx 0.09$.

Decreasing returns to R&D $\{\zeta\}$. From Equation (1.12), the relationship between the number of inventors that file patents and the number of a firm's hired inventors is log-linear, such that $\log\left(n_o^{\omega}\right) = \zeta \log\left(R_o^{\omega}\right)$. When taking this expression to the data, I run the following regression:

$$\log\left(n_{o,t}^{\omega}\right) = \iota + \iota_{\omega} + \iota_{o,t} + \zeta \log\left(R_{o,t}^{\omega}\right) + \epsilon_{o,t}^{\omega}, \tag{1.35}$$

where ι_{ω} are firm fixed effects, $\iota_{o,t}$ are location/period fixed effects, and $\epsilon_{o,t}^{\omega}$ are i.i.d shocks. Using data on the number of firms' inventors $R_{o,t}^{\omega}$ and the number of firms' inventors that generated an idea $n_{o,t}^{\omega}$, I regress ζ directly from the data. To keep the estimation consistent with the reduced-form estimates, I consider 10-year periods. Column (3) of Table 3 shows the main estimate of $\zeta = 0.65$, which confirms the existence of decreasing returns to R&D in the data.

This estimate is around the upper bound of previously estimated values in the literature between 0.1 and 0.6 (Kortum 1993). The main reason is that I estimate long run decreasing returns to R&D since I consider 10-year periods in the estimation. This contrasts with previous estimates that use yearly or cross-sectional variation. To see this, Table A6 contains additional estimations of ζ when considering 5-year periods. Indeed, the value of this estimate goes down to $\zeta = 0.568$, which suggests stronger decreasing returns to R&D in the short run.

Table 3: Estimation of decreasing returns to R&D

	(1)	(2)	(3)
$\log\left(R_{o,t}^{\omega}\right)$	0.718	0.704	0.65
	(0.009)	(0.0096)	(0.0103)
$\iota_{o,t}$		✓	✓
ι_{ω}			✓
N	$49,297$	$49,297$	$25,010$
R^2	0.72	0.812	0.904

Notes: In this table I report estimates for decreasing returns to R&D from Equation (1.35). The dependent variable $\log\left(n_{o,t}^{\omega}\right)$ is the number of firm's inventors that filed a patent. Each column is an specification with different combinations of fixed effects. The fixed effects included in each specification are determined by rows $4-5$. Row 2 contains the estimates for ζ, and row 3 contain standard errors, which are clustered at the o,t level. Rows $6-7$ contain the number of observations and goodness of fit in each specification, respectively.

Migration costs $\{\mu_{od}^{n}\}$. For each agent type $n = \{L, R\}$, I parametrize migration costs as an exponential function of geographic distance between every location pair $\mu_{od}^{n} = \rho_0^n dist_{od}^{\rho_1^n} \exp\left(-\frac{\epsilon_{od}^{n}}{\kappa}\right)$, where $\{\rho_0^n\}$ are intercepts that determines the overall level of internal migration, $\{\rho_1^n\}$ are the elasticities of migration costs to distance, and ϵ_{od}^{n} are i.i.d.

shocks. To keep the estimation consistent with the reduced-form estimates, I consider 10-year periods. I calibrate $\{\rho_0^n\}$ by targeting the 10-year average migration rates for workers and inventors of 24.99% and 26.38%, respectively. The calibrated values are $\left\{\rho_0^L, \rho_0^R\right\} = \{1.361, 1.354\}$. To estimate $\{\rho_1^n\}$, the location choice problem of the model yields migration gravity equations for both workers and inventors:

$$\log\left(\eta_{od,t}^n\right) = \iota + \iota_{o,t} + \iota_{d,t} - \kappa\rho_1^n \log\left(dist_{od}\right) + \epsilon_{od,t}^n, n = \{L, R\}. \tag{1.36}$$

The gravity equation in (1.36) states that, conditional on origin/time and destination/time fixed effects $\{\iota_{o,t}, \iota_{d,t}\}$, data on geographic distance between locations, and the spatial labor supply elasticity κ, the migration elasticities to trade costs $\{\rho_1^n\}$ are identified. Since migration shares report values of zero, I estimate these elasticities through Poisson Pseudo Maximum Likelihood (PPML) estimation. From columns (2) and (4) in Table 4, I consider $\left\{\rho_1^L, \rho_1^R\right\} = \{0.591, 0.602\}$. These values are very close to the median value of migration elasticities estimated by Allen and Donaldson (2020), who also estimate them considering 10-year periods. Intuitively, Table A7 shows that the value of these elasticities go up to $\left\{\rho_1^L, \rho_1^R\right\} = \{0.65, 0.651\}$ when considering 5-year periods, which reflect higher barriers to move in the shorter run.

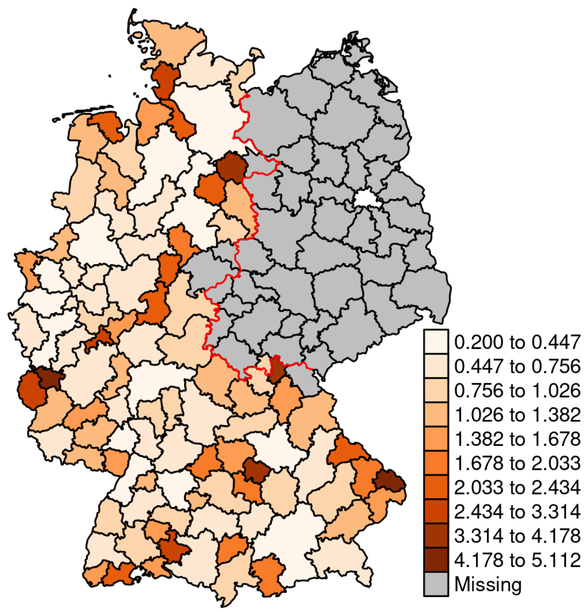
Table 4: Estimation of migration costs

	$n = R$		$n = L$	
	OLS	$PPML$	OLS	$PPML$
$\log\left(dist_{od}\right)$	-1.001	-1.254	-1.063	-1.277
	(0.014)	(0.018)	(0.020)	(0.016)
ρ_1^n	0.472	0.591	0.501	0.602
R^2	0.812	·	0.839	·
N	8,336	21,632	18,381	21,632

Notes: In this table I report migration cost elasticities from Equation (1.36). Columns $2-3$ are the regressions for inventors, where column 2 are OLS estimates, and column 3 are PPML estimates. Columns $4-5$ are the regressions for workers, where column 4 are OLS estimates, and column 5 are PPML estimates. For OLS estimates, the dependent variable is measured as $\log\left(\eta_{od,t}^n\right)$ is the log of the share of inventors or workers from o that moved to d during a given period. Row 3 is the estimate associated to $\log\left(dist_{od}\right)$, where $dist_{od}$ is the Euclidean distance in miles from o to d. Row 4 are standard errors two-way clustered at the o, t and d, t level. Row 5 is the implied migration elasticity from the estimates from row 3 given $\kappa = 2.12$. Rows $6-7$ contain the goodness of fit and number of observations in each specification, respectively.

Fundamental location productivity $\{\mathcal{A}_o\}_{\forall o \in \mathcal{S}}$. I follow Redding (2016) to recover unobserved fundamental location productivities $\{\mathcal{A}_o\}_{\forall o \in \mathcal{S}}$ through model inversion. Given values for parameters $\{\sigma, \widetilde{\gamma}\}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, and data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o \in \mathcal{S}}$, there is a unique set of values for fundamental location productivities $\{\mathcal{A}_o\}_{\forall o \in \mathcal{S}}$ that is consistent with the data. Since the model is static, I use data on wages and population from 2014 to denote West Germany' steady-state equilibrium. To recover these fundamentals, I solve a fixed point algorithm on the system of excess demand functions implied by Equations (1.32). In Figure 1 I show the spatial distribution of these fundamentals. As expected, to rationalize the presence of production and innovation in less-dense locations, these locations must report higher fundamental levels of productivity.

Figure 1: Fundamental location productivity



Notes: This figure shows the spatial distribution of fundamental location productivities $\{\mathcal{A}_o\}$ in West Germany. A darker (lighter) orange color denotes a higher (lower) productivity. All these values are normalized by their corresponding geometric mean.

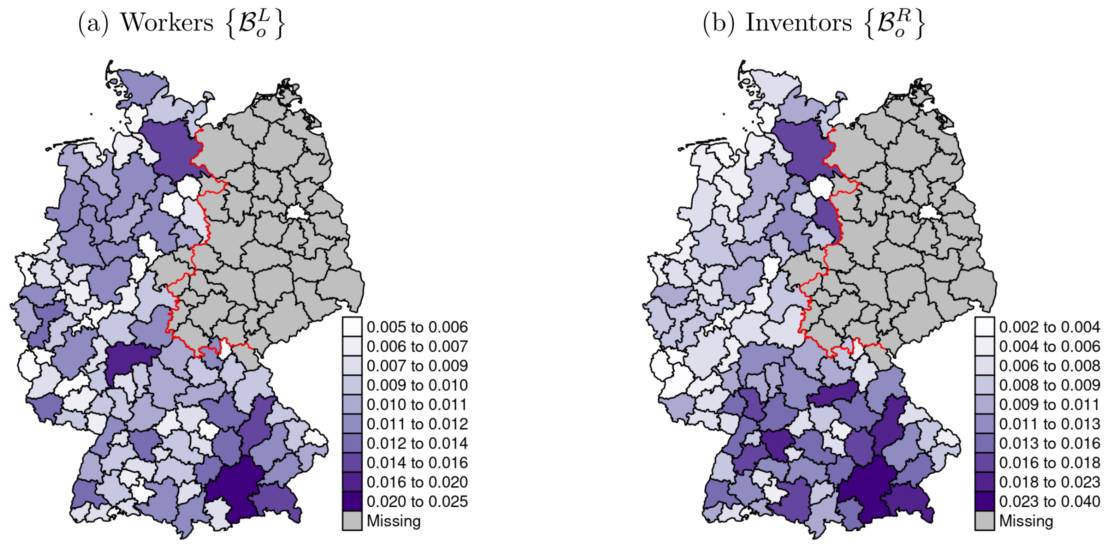
Fundamental location amenities $\{\mathcal{B}_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$. I follow Redding (2016) to recover location fundamental amenities for both workers and inventors $\{\mathcal{B}_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$ through model inversion. Given values for parameters $\{\sigma, \kappa, \widetilde{\gamma}\}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, migration costs $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, fundamental location productivities $\{\mathcal{A}_o\}_{\forall o \in \mathcal{S}}$, an initial distribution of workers and inventors across locations $\left\{\overline{L}_o, \overline{R}_o\right\}_{\forall o \in \mathcal{S}}$, and data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o \in \mathcal{S}}$, there is a unique set of values for fundamental location amenities $\{\mathcal{B}_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$ that is consistent with the data. Since the model is static, I use data on wages and population from 2014 to denote West Germany' steady-state equilibrium. The initial distribution $\left\{\overline{L}_o, \overline{R}_o\right\}_{\forall o \in \mathcal{S}}$ is from 1980 and they are scaled such that the total number of workers and inventors in West Germany is the same for 2014. To recover these fundamentals, I solve a fixed point algorithm on the system of excess demand functions implied by Equations (1.29) and (1.27).

In Figure 2 I show the spatial distribution of these fundamentals. As reflected by the spatial distribution of both workers and inventors, locations like Munich, Stuttgart, and Hamburg reflect the highest levels of amenities. More importantly, inventors exhibit higher levels of location amenities in the south of West Germany than workers, which reflects their higher level of spatial concentration in the data.

Fundamental location amenities $\{\mathcal{B}_o^n\}_{\forall o\in\mathcal{S}}^{n=\{L,R\}}$. I follow Redding (2016) to recover location fundamental amenities for both workers and inventors $\{\mathcal{B}_o^n\}_{\forall o\in\mathcal{S}}^{n=\{L,R\}}$ through model inversion. Given values for parameters $\{\sigma,\kappa,\widetilde{\gamma}\}$, trade costs $\{\tau_{od}\}_{\forall o,d\in\mathcal{S},\mathcal{S}}$, migration costs $\{\mu_{od}^n\}_{\forall o,d\in\mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, fundamental location productivities $\{\mathcal{A}_o\}_{\forall o\in\mathcal{S}}$, an initial distribution of workers and inventors across locations $\left\{\overline{L}_o,\overline{R}_o\right\}_{\forall o\in\mathcal{S}}$, and data on wages and population $\left\{w_o^L,w_o^R,L_o,R_o\right\}_{\forall o\in\mathcal{S}}$, there is a unique set of values for fundamental location amenities $\{\mathcal{B}_o^n\}_{\forall o\in\mathcal{S}}^{n=\{L,R\}}$ that is consistent with the data. Since the model is static, I use data on wages and population from 2014 to denote West Germany' steady-state equilibrium. The initial distribution $\left\{\overline{L}_o,\overline{R}_o\right\}_{\forall o\in\mathcal{S}}$ is from 1980 and they are scaled such that the total number of workers and inventors in West Germany is the same for 2014. To recover these fundamentals, I solve a fixed point algorithm
on the system of excess demand functions implied by Equations (1.29) and (1.27).

In Figure 2 I show the spatial distribution of these fundamentals. As reflected by the spatial distribution of both workers and inventors, locations like Munich, Stuttgart, and Hamburg reflect the highest levels of amenities. More importantly, inventors exhibit higher levels of location amenities in the south of West Germany than workers, which reflects their higher level of spatial concentration in the data.

Figure 2: Fundamental location amenities



Notes: This figure shows the spatial distribution of fundamental location amenities $\{\mathcal{B}_o^L, \mathcal{B}_o^R\}$ in West Germany. A darker (lighter) purple color denotes a higher (lower) productivity. All these values are normalized by their corresponding arithmetic mean.

Trade costs $\{\tau_{od}\}$**.** I parametrize trade costs as an exponential function of geographic distance between every location pair $\tau_{od} = \xi_0 dist_{od}^{\xi_1}$, where ξ_0 is an intercept that determines the overall level of internal trade, and ξ_1 is the elasticity of trade costs to distance. Following Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016), I calibrate ξ_0 to target a 50% share of total intra-regional trade. For the elasticity of trade costs to distance, I follow Krebs and Pflüger (2021) and set $\xi_1 = \frac{1.56}{\sigma-1}$. I use this value instead of the one from Monte, Redding, and Rossi-Hansberg (2018) since the former is based on internal trade data for Germany.

Remaining parameters $\{\alpha, \kappa, \sigma\}$**.** The remaining parameters are the dispersion of productivity of ideas α, the spatial labor supply elasticity κ, and the elasticity of substitution across intermediate inputs σ. Regardless of the microfoundation for firms' R&D, a value of α is necessary to obtain values for the constant ψ from Equation (1.15). Following the process for the generation of ideas from Appendix A.4.1, α is the Pareto shape parameter for the productivity of ideas. I run a parametric fit on the number of 5-year forward citations and set $\alpha = 1.5$. This value is similar to previous Pareto parametric fits for the number of forward citations (Silverberg and Verspagen 2007). For the migration elasticity κ, I follow Peters (2022) and set $\kappa = 2.12$ since it is estimated for the German context. Finally, I follow Broda and Weinstein (2006) and set $\sigma = 2.5$, which is the median elasticity for industrial sectors, which corresponds to the German context between 1980 and the 2000s.

Table 5: Summary of calibration

	Description	Value	Identification/Moments
		Innovation	
$\widetilde{\gamma}$	Spatial knowledge spillovers in R&D	$\widetilde{\gamma} = (0.409)\,(0.22)$	0.409 : IV estimate, Table 2, column 3 0.22 : Lanjouw and Schankerman (2004), Table 2, column 8
ζ	Decreasing returns to R&D	0.65	OLS estimate, Table 3, column 3
α	Idea productivity dispersion	1.5	Pareto parametric fit
		Migration	
$\{\rho_0^R, \rho_0^L\}$	Migration costs, intercepts	$\rho_0^R = 1.354$ $\rho_0^L = 1.361$	26.38% migration rate of inventors 24.99% migration rate of workers
$\{\rho_1^R, \rho_1^L\}$	Migration costs, elasticities	$\rho_1^R = \frac{1.254}{\kappa}$ $\rho_1^L = \frac{1.277}{\kappa}$	Gravity estimates
κ	Migration elasticity	2.12	Peters (2022), Table 9
		Location fundamentals	
$\mathcal{A}_o$	Location productivites		Recovered, Equation (1.32)
$\{\mathcal{B}_o^R, \mathcal{B}_o^L\}$	Location amenities		Recovered, Equations (1.27) and (1.29)
		Trade	
ξ_0	Trade costs, intercept	0.17	50% intra-trade shares (Ramondo, Rodríguez-Clare, and Saborío-Rodríguez 2016)
ξ_1	Trade costs, elasticity	$\frac{1.56}{\sigma-1}$	Krebs and Pflüger (2021)
σ	Elasticity of substitution	2.5	Broda and Weinstein (2006), Table 5

Notes: This table summarizes the calibration of the model parameters. The first column shows the parameter of interest, the second column provides a short description, the third column reports the calibrated value, and the fourth column briefly describes the identification strategy.

1.5 Counterfactuals

In this section, I use the calibrated model to conduct two set of counterfactuals. First, I quantify the effect of reducing inventor migration costs by 25%. Second, I quantify the effect of the 2020 German R&D Tax Allowance Act which implemented a 25% subsidy for firms' R&D expenditure. In each counterfactual, I study the effect of these policies on aggregate productivity, and explore how these effects depend on spatial knowledge spillovers in R&D.

1.5.1 Reducing inventor migration costs

The model predicts that reducing inventor migration costs μ_{od}^R by 25% leads to an increase of aggregate productivity of 5.87%. This figure is comparable to Bryan and Morten (2019) who find that a 30% proportional reduction of both μ_{od}^L and μ_{od}^R lead to a 7% increase of aggregate output. Nevertheless, it is surprising that a similar reduction of μ_{od}^R can lead to comparable increases in aggregate productivity or output since inventors comprise a small share of the population. In that sense, fostering the mobility of the agents behind R&D could be a cost-effective way to promote economic activity.

Additionally, since the number of inventors in the economy is finite, this policy could generate heterogeneous effects across locations. In Figure 3 I analyze the productivity gains of each location, where these gains are measured as

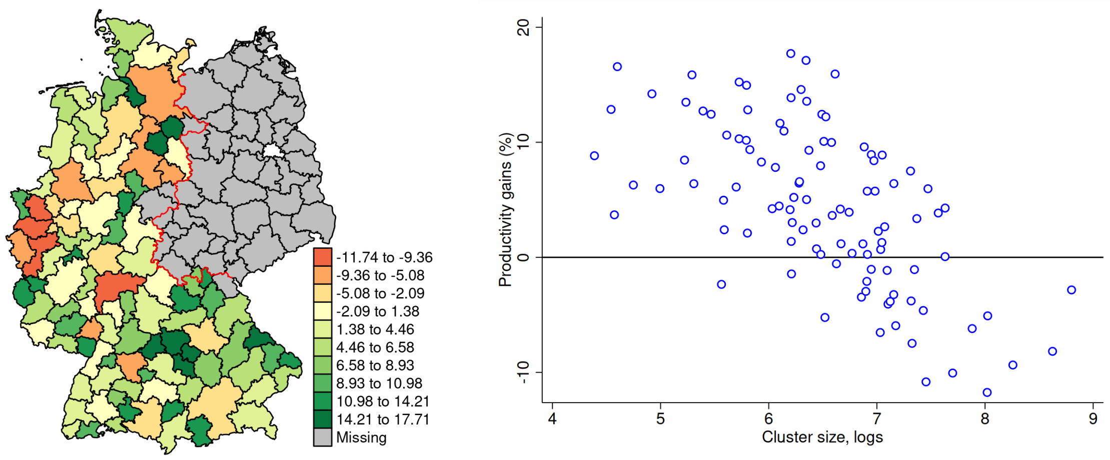
$$g_o^A = \left(\frac{A_o^{policy} - A_o^{baseline}}{A_o^{baseline}} \right) \times 100\%,$$

where A_o^{policy} is the productivity of location o under the reduction of migration costs for inventors, and $A_o^{baseline}$ is productivity of location o at the baseline scenario. The left panel of the figure shows a map with the location productivity gains of implementing the policy. The map shows that the policy indeed generates large heterogeneous effects across locations. At the upper tail of the distribution, there are locations that increase their productivity by around $14\% - 17\%$. In contrast, at the lower tail, some locations exhibit lower productivity by around $9\% - 11\%$.

On the right panel of Figure 3, we observe that reducing μ_{od}^R exhibits an equalizing effect—larger clusters at the baseline exhibited decreases in productivity, while smaller clusters gained productivity—That is, after facilitating the spatial mobility of inventors, they prefer to move towards smaller clusters. For example, after the policy, a large share of inventors from Munich and Stuttgart moved to contiguous locations, so they exhibited the largest productivity gains. Inventors move towards smaller clusters after the policy because they can earn higher real wages.

Figure 3: Reduction of inventor migration costs by 25%

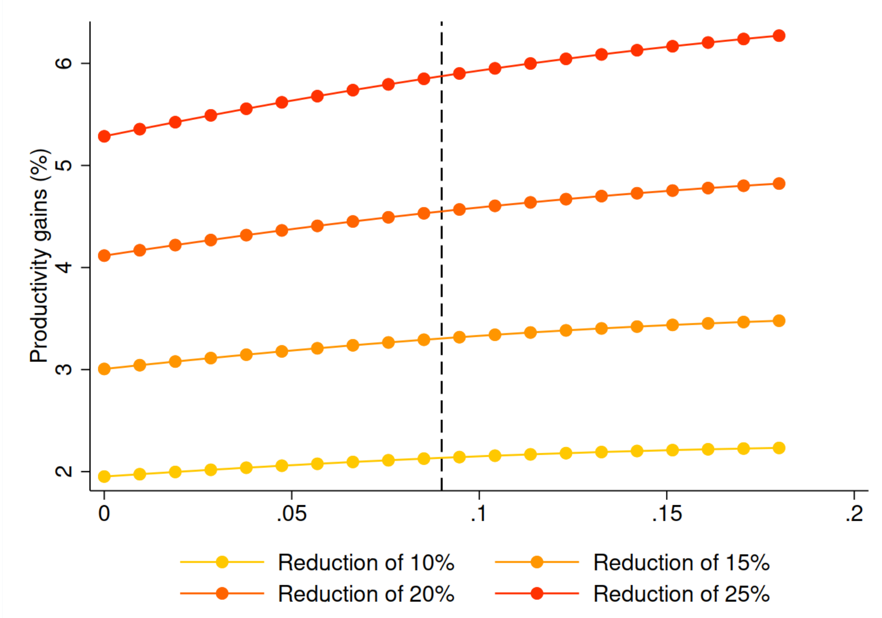
(a) Productivity gains, map (b) Productivity gains VS cluster size in logs



Notes: This figure is comprised by two panels. I measure productivity gains as $g_o^A = \left(\frac{A_o^{policy} - A_o^{baseline}}{A_o^{baseline}}\right) \times 100\%$, where A_o^{policy} is aggregate productivity of location o under the 25% reduction of μ_{od}^R, and $A_o^{baseline}$ is aggregate productivity of location o at the baseline scenario. On the left panel, I color each location in West Germany according to their value of g_o^A. On the right panel, I compare g_o^A with the number of inventors in each location at the baseline $R_o^{baseline}$.

I now explore how spatial knowledge spillovers in R&D influence the effect of reducing μ_{od}^R on aggregate productivity. In Figure 4, intuitively, we observe that reducing μ_{od}^R unambiguously increases aggregate productivity. For example, consider the dashed vertical line for the calibrated value of $\widetilde{\gamma} = 0.09$. Then, larger reductions of μ_{od}^R exhibit larger productivity gains. More interestingly, the effect of reducing μ_{od}^R on aggregate productivity exhibits complementarity with the value of $\widetilde{\gamma}$. Considering a reduction of 10% (the yellow line), going from scenario of no spillovers ($\widetilde{\gamma} = 0$) to doubling the spillovers ($\widetilde{\gamma} = 0.09 \times 2$) generates additional 28*pp* productivity gains. In contrast, considering a reduction of 25% (the red line), the same exercise leads to additional 98*pp* productivity gains. This highlights the importance of implementing policies that foster both the mobility of inventors and spatial knowledge spillovers in R&D to promote economic activity.

Figure 4: Productivity gains VS $\widetilde{\gamma}$, by reduction of μ_{od}^R



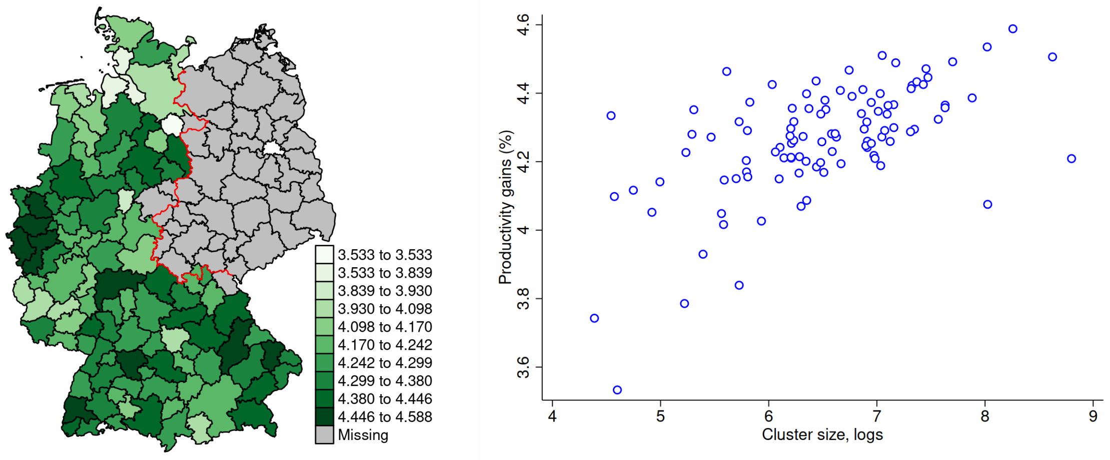
Notes: This figure shows the relationship between productivity gains g^A of reducing μ_{od}^R VS spatial knowledge spillovers $\widetilde{\gamma}$. I measure productivity gains as $g^A = \left(\frac{A_o^{policy} - A_o^{baseline}}{A_o^{baseline}}\right) \times 100\%$, where A_o^{policy} is aggregate productivity of location o under the reduction of μ_{od}^R, and $A_o^{baseline}$ is aggregate productivity of location o at the baseline scenario. The horizontal axis exhibits different values of $\widetilde{\gamma}$, where the vertical dashed line indicates the calibrated value of $\widetilde{\gamma} = 0.09$. I plot four scenarios ranging from a reduction of μ_{od}^R of 10% (yellow) to a reduction of 25% (red).

1.5.2 2020 German R&D Tax Allowance Act

In this section I evaluate the 2020 German R&D Tax Allowance Act, which introduced a R&D tax incentive scheme as from January 1st 2020. Under this scheme, firms were entitled to receive funding for their R&D activities. In particular, this scheme provides a 25% subsidy for in-house R&D activities regardless of firm characteristics (Deloitte 2020). The model predicts that implementing an R&D subsidy $s_o = 25\%$ increases aggregate productivity by 4.27%. In Figure 5 we see that the implications for heterogeneity across locations is drastically different from the policy of reducing inventor migration costs. First, the left panel shows that the subsidy increases aggregate productivity everywhere, ranging from 3.5% to 4.5% gains. Second, the right panel shows that larger clusters are the ones that exhibited larger productivity gains since they increased even more in size, so the subsidy increased the spatial concentration of inventors.

Figure 5: Subsidy for firms' R&D expenditure by 25%

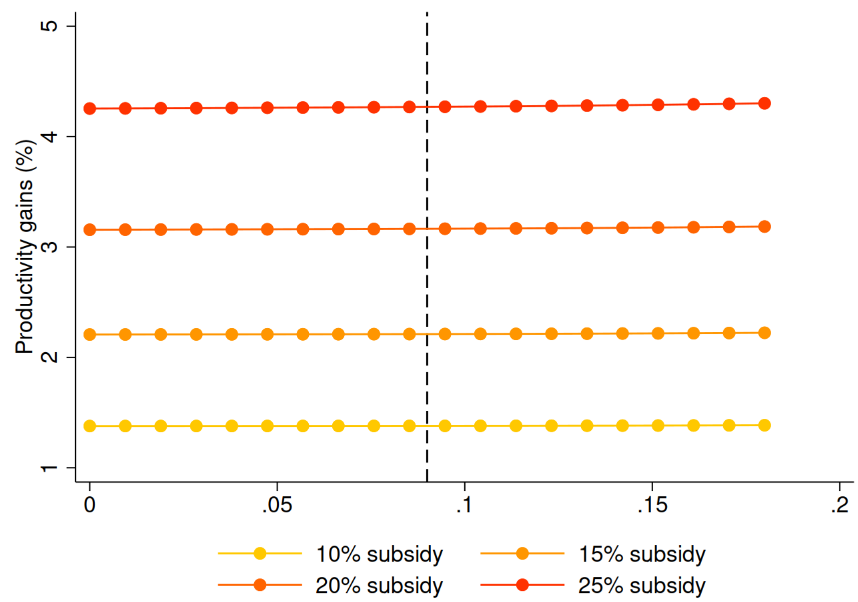
(a) Productivity gains, map (b) Productivity gains VS cluster size in logs



Notes: This figure is comprised by two panels. I measure productivity gains as $g_o^A = \left(\frac{A_o^{policy} - A_o^{baseline}}{A_o^{baseline}}\right) \times 100\%$, where A_o^{policy} is aggregate productivity of location o under the R&D subsidy $s_o = 25\%$, and $A_o^{baseline}$ is aggregate productivity of location o at the baseline scenario. On the left panel, I color each location in West Germany according to their value of g_o^A. On the right panel, I compare g_o^A with the number of inventors in each location at the baseline $R_o^{baseline}$.

Now, I now explore how spatial knowledge spillovers in R&D influence the effect of implementing R&D subsidies on aggregate productivity. In Figure 6, intuitively, we observe that implemeting an R&D subsidy unambiguously increases aggregate productivity. In contrast with the policy of reducing inventor migration costs, the effect of implementing the subsidy on aggregate productivity exhibits weaker complementarity with the value of $\widetilde{\gamma}$. Considering a subsidy of 10% (the yellow line), going from scenario of no spillovers ($\widetilde{\gamma} = 0$) to doubling the spillovers ($\widetilde{\gamma} = 0.09 \times 2$) generates additional 0.7pp productivity gains. In contrast, considering a subsidy of 25% (the red line), the same exercise leads to additional 4.7pp productivity gains.

Figure 6: Productivity gains VS $\widetilde{\gamma}$, by value of s_o



Notes: This figure shows the relationship between productivity gains g^A of implementing an R&D subsidy $s_o = 25\%$ VS spatial knowledge spillovers $\widetilde{\gamma}$. I measure productivity gains as $g^A = \left(\frac{A_o^{policy} - A_o^{baseline}}{A_o^{baseline}}\right) \times 100\%$, where A_o^{policy} is aggregate productivity of location o under the R&D subsidy, and $A_o^{baseline}$ is aggregate productivity of location o at the baseline scenario. The horizontal axis exhibits different values of $\widetilde{\gamma}$, where the vertical dashed line indicates the calibrated value of $\widetilde{\gamma} = (0.409)(0.22) = 0.09$. I plot four scenarios ranging from $s_o = 10\%$ (yellow) to $s_o = 25\%$ (red).

1.6 Conclusions

In this paper I quantify the importance of spatial knowledge spillovers in R&D for aggregate productivity. I causally estimate these spillovers by exploiting the historical episode of the arrival of East German inventors across West Germany after the Reunification of Germany. I then embed these spillovers into a spatial model of innovation, and use the model to quantify the importance of these spillovers when implementing policies that promote R&D activities for aggregate productivity. I show that reducing migration costs for inventors and subsidies to firms' R&D activities can substantially increase aggregate productivity, and spatial knowledge spillovers in R&D is crucial for the effectiveness of these policies.

This paper have abstracted from other different channels that could also contribute aggregate productivity. First, occupational choice between workers and inventors, or firm selection into R&D through firm heterogeneity could amplify the effect of policies due to entry of agents into innovation. Second, inter-temporal knowledge spillovers could be introduced in the model to quantify the role of spatial knowledge spillovers in R&D and R&D policies for long-run growth. Finally, new micro-data on inventors also allows to account for the importance of firm-level spillovers and the rise of teams. The model is flexible enough to easily introduce these mechanisms, I leave these for future research.

Chapter II. Cultural Proximity and Production Networks

with Gaurav Khanna and Hiroshi Toma

2.1 Introduction

Non-economic forces, such as *culture*—, religion, language, values, etc.—drive economic outcomes. The role of culture on agent behavior has been well documented in entrepreneurship, loan access, labor markets, marriage, and international trade (Bandyopadhyay, Coughlin, and Wall 2008; Fisman, Paravisini, and Vig 2017; Goraya 2022; Guiso, Sapienza, and Zingales 2009; Hasanbasri 2019; Macchiavello and Morjaria 2015; Munshi and Rosenzweig 2016; Rauch 1996; Rauch and Casella 2003; Rauch and Trindade 2002; Schoar, Iyer, and Kumar 2008; Startz 2016; Zhou 1996). At the same time, recent evidence increasingly shows how inter-firm trade and production networks have important aggregate implications for economic development and welfare (Antras, Fort, and Tintelnot 2017; Bernard et al. 2009; Bernard, Moxnes, and Saito 2019; Bernard and Moxnes 2018; Bernard et al. 2022; Dhyne, Kikkawa, and Magerman 2022; Eaton, Kortum, and Kramarz 2011; Eaton et al. 2016; Huneeus 2018; Lim 2018; Munshi and Rosenzweig 2016; Oberfield 2018; Taschereau-Dumouchel 2020). Despite their parallel importance, the mechanisms by which cultural proximity shapes production networks and their aggregate implications remain less understood. Understanding how and why cultural proximity affects firm linkages and trade, potentially allows policy-makers to better leverage social inclusion programs and foster economic development. In this paper, we examine how cultural proximity determines connections and trade within production networks, and quantify the implications of cultural links for welfare and productivity.

We first provide empirical evidence on the role of cultural proximity in inter-firm trade and the formation of production networks. To do this, we leverage a unique dataset of firm-to-firm transactions from a large Indian state, along with data on firm owners names and their cultural proximity derived from India's caste and religious system. We report three new stylized facts. First, culturally closer firms report higher sales between them: the higher the cultural proximity, the higher the trade on the intensive margin. Second, culturally closer firms are more likely to ever trade with each other. This means the higher the cultural proximity, the higher the trade on the extensive margin as well. Third, firms that are culturally further apart report higher unit prices in their transactions. All these results are robust to an array of high-dimensional fixed effects, including seller and buyer fixed effects, origin-by-destination fixed effects (and for specifications with product and time, seller-by-product, and product-by-month fixed effects).

We then turn to explore the importance of contract enforcement. First, we show suggestive evidence that the effect we find of cultural proximity on trade is driven by differentiated goods, which often rely on either formal or informal contract enforcement (Nunn 2007; Rauch 1999). Indeed, we find that differentiated goods, are more likely to be produced in and bought by firms that are located in districts with higher contract enforcement (as proxied by court delays). We understand these findings as evidence that cultural proximity relates to contract enforcement and trust (Munshi 2014; Munshi 2019).[4]

Differentiated goods do not trade in exchanges and are not homogeneous, but are branded and specific to certain producing firms. In a country with market imperfections as India, firms can easily renege on their commitments. Suppliers and buyers in differentiated goods markets are not easily replaceable. In such cases, trade will increase when firms trust and know each other, that is, when they are culturally close.

We further find that the more varieties a firm sells or buys, the more the trade intensity is affected by social proximity. We posit that the larger the amount of different varieties a firm sells or buys, the more firms it has to negotiate with, which increases the contracting frictions it faces. Then, in order to minimize the contracting frictions they face, firms will rely more on trading with culturally closer firms they trust.

To analyze whether our results are caused by vertical social hierarchies and discrimination across cultural groups, we study asymmetric effects in those transactions where one firm is placed higher than the other based on the caste-based hierarchy, allowing us to test for preference-based discrimination across the social hierarchy. We do not find much evidence that hierarchies (and preference-based discrimination) across social groups matter for our social proximity results. In other tests, we find our results are less likely to be driven by firms sharing the same language or specialization in the production of certain goods.

Encouraged by these stylized facts, we build a quantitative general equilibrium model of firm-to-firm trade and cultural proximity. Firms produce goods by combining labor and intermediate inputs in a CES fashion. Firms sell their goods to a household as final goods

[4] Munshi (2019) uses survey data to show that Indians trust people from their caste. He also gives an example on how the Indian diamond industry relies on community networking because of the deficient contract enforcement.

and to other firms as intermediates. Firms engage in monopolistic competition, charging a constant markup on top of their marginal costs. Importantly, we introduce our measure of cultural proximity as a wedge that affects both trade and matching costs.

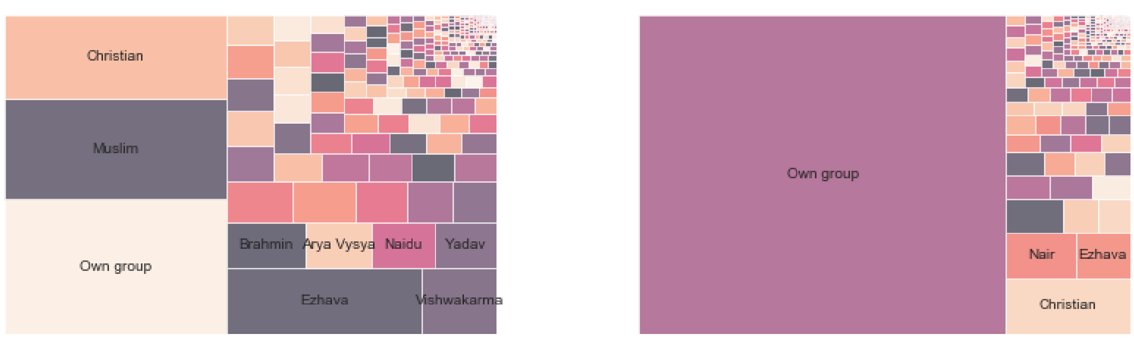
The model derives equations that precisely match their empirical counterparts in the previous section. We use these equations to estimate the key parameters of the model: the semi-elasticity of the trade cost to cultural proximity and the semi-elasticity of matching cost to cultural proximity. Our model allows us to estimate both of these parameters externally. In line with our stylized facts, we find a negative semi-elasticity of both the intensive and extensive margin of trade to cultural proximity. This implies the closer two firms are in cultural terms, the lower the trade and matching costs are. Therefore, the higher the cultural proximity for a pair of firms, the higher the trade is on both the intensive and extensive margins, and the lower the prices charged.

We use the model and estimated parameters to quantify the implications for welfare and other aggregate outcomes of implementing different policies. First, we evaluate the effects of social mixing/inclusion (i.e. firms become culturally the closest possible) and social isolation policies (i.e. firms become culturally the furthest possible). Second, we study the effects of a policy that reduces contracting frictions, such that firms rely less on cultural proximity when trading (i.e. trade and matching costs become less sensitive to cultural proximity). We find that welfare increases by 1.76 percent under a diversity-friendly social inclusion policy. In contrast, welfare falls by 1.45 percent when we evaluate the effects of social isolation or exclusion. Finally, we show that policies that reduce contracting frictions raise welfare by 0.87 percent by reducing the reliance of trade on cultural links.

Figure 7: Probability-weighted sales decomposition of largest cultural groups

(a) Largest Hindu group: Nair

(b) Largest non-Hindu group: Muslims



Notes: Figure shows the decomposition across buyers for the largest Hindu and non-Hindu cultural groups measured by probability-weighted sales. The Nair and Muslims accounted for 4.88 and 11.83 percent of total probability-weighted sales, respectively.

The analysis of cultural proximity is especially relevant for developing countries, where agents face several contracting frictions and, consequently, rely more on non-economic forces (Boehm and Oberfield 2020; Munshi and Rosenzweig 2016; Munshi 2019). In particular, India has a society that follows the parameters of a caste system, which also intertwines with the different religious groups.[5] In this case, cultural proximity naturally arises as a product of the inherent hierarchical structure of the caste system and the different religions. Related to this, Figure 7 shows an example of how trade between cultural groups occurs, in a selected subset of our data. We can see that there are cultural groups that are bound to trade more or less with other cultural groups. We thus ask whether cultural proximity, measured as the cultural group-based distance between firms, can determine trade.

This paper contributes to two strands of the literature. First, the paper contributes to the role of cultural proximity on economic outcomes such as trade (Bandyopadhyay, Coughlin, and Wall 2008; Guiso, Sapienza, and Zingales 2009; Macchiavello and Morjaria 2015; Rauch 1996; Rauch and Casella 2003; Rauch and Trindade 2002; Richman 2006; Schoar, Iyer, and Kumar 2008; Startz 2016; Zhou 1996), entrepreneurship (Goraya 2022), finance (Fisman, Paravisini, and Vig 2017), and labor markets (Munshi and Rosenzweig 2016; Hasanbasri 2019). Second, it contributes to work on production networks (Antras, Fort, and Tintelnot 2017; Bernard et al. 2009; Bernard, Van Beveren, and Vandenbussche 2014; Bernard, Moxnes, and Saito 2019; Bernard and Moxnes 2018; Bernard et al. 2022; Dhyne, Kikkawa, and Magerman 2022; Eaton, Kortum, and Kramarz 2011; Eaton et al. 2016; Eaton, Kortum, and Kramarz 2022; Huneeus 2018; Lim 2018; Oberfield 2018; Taschereau-Dumouchel 2020). We merge these two separate strands of the literature by providing both evidence and theory on how cultural proximity between firms can shape inter-firm trade, and what this implies for aggregate welfare. The uniqueness of our data in terms of measuring firm-to-firm transactions and the cultural group of owners, in combination with substantial variation across cultural groups, allow us to answer how cultural proximity shapes linkages and trade across the production network.

The rest of the paper is structured as follows. In Section 2.2 we provide a brief review of the caste system in India, describe our new datasets and explain how we construct firm-level trade and cultural proximity variables. In Section 2.3 we report our stylized facts. In Section 2.4 we describe the model. In Section 2.5 we explain how we estimate the key parameters of the model. In Section 2.6 we analyze counterfactual scenarios. Section 2.7 concludes.

[5] In this paper, we consider the caste system and the religious groups as a proxy for cultural groups. There is a large historical legacy for the caste system to be considered as a device for discrimination, which we consider. Even though there is an active agenda of the government to implement policies that hinder caste-based discrimination, it is still used by Indians as a way to determine how similar individuals are between them.

2.2 Background, data and construction of variables

2.2.1 Caste and Religion in India

India has a society that is heavily influenced by the parameters of a caste system: a hierarchical system that has prevailed in the country since around 1,500 BC and that still rules its economy. According to this classification, people are classified across four possible groups called *Varnas*. From the most to the least privileged in hierarchical order, the four Varnas are *Brahmins*, *Kshatriyas*, *Vaishyas*, and *Shudras*. The Brahmins have historically enjoyed the most privileges, and are traditionally comprised of priests and teachers. The Kshatriyas are next in the hierarchy, usually associated with a lineage of warriors. The Vaishyas are third and are related to businessmen such as farmers, traders, among others. Finally, the Shudras are the most discriminated against and are the caste formed to be the labor class.

At the same time, Varnas are comprised by sub-groups called *Jatis* that were determined by factors such as occupation, geography, tribes, or language. In that sense, using Jatis as castes are appropriate for studying economic networks (Munshi 2019), and from here on we use the notion of Jatis when referring to castes.

We also consider religious groups to define other cultural groups. The caste system is inherently based on Hindu religion, the predominant religion in India. While there are other religions in India which do not follow the caste system, they do relate to it: the other non-Hindu religions work as cultural groups of their own. We leverage information on firm owners belonging to both caste and religious groups to construct our measure of cultural proximity.

2.2.2 Data

Firm-to-firm trade. We leverage a firm-to-firm trade dataset for a large Indian state provided by the state's corresponding tax authority.[6] We use daily transactions data from January 2019 to December 2019, as long as at least one node of the transaction (either origin or destination) was in the state. This data exists due to the creation of the E-Way bill system in India on April 2018, where firms register the movements of goods online for tax purposes. This is a major advantage over traditional datasets collected for tax purposes in developing countries since the E-Way bill system was created with the purpose of significantly increasing tax compliance.[7]

This data is provided by the tax authority of a large Indian state with a diversified production structure, roughly 50 percent urbanization rates, and high levels of population

[6]While we use the term 'firm' in most parts of the paper, these data are actually at the more granular establishment level.

[7]For more details about the new E-Way bill system, see `https://docs.ewaybillgst.gov.in/`

density. To compare its size in terms of standard firm-to-firm transaction datasets, the population of this Indian state is roughly three times the population of Belgium, seven times the population of Costa Rica, and double the population of Chile. In addition, we can uniquely measure product-specific prices for each transaction, along with the usual measures of total value traded.

Each transaction reports a unique tax code identifier for both selling and buying firm. We use these identifiers to merge this data with other firm-level datasets. We also have information on all the items contained within the transaction, the value of the transaction, the 6-digit HS code of the traded items, the quantity of each item and the units of the quantity is measured in. Since the data report both value and quantity of traded items, we construct unit values for each transaction. Each transaction also reports the pincode (zip code) location of both selling and buying firms. By law, any person dealing with the supply of goods and services whose transaction value exceeds 50,000 Rs (700 USD) must generate E-way bills. Transactions that have values lower than 700 USD can also be registered but it is not mandatory. There are three types of recorded transactions: (i) within-state trade, (ii) across-states trade, and (iii) international trade. For the purpose of this paper, we ignore international trade.

Firm owner names. The information about the name of the firm owners comes from two different sources. The first source is also provided by the tax authority of the Indian state, which is a set of firm-level characteristics for firms registered within our large Indian state. Among these variables, we are provided with the name of the owner and/or of representatives of the firm.

To obtain firm-level characteristics of firms not registered in this state, we scrape the website *IndiaMART*,[8] the largest e-commerce platform for business-to-business (B2B) transactions in India. The website is comprised of firms of all sizes. By 2019, the website registered around 5-6 million sellers scattered all around India. Most importantly, this platform provides the name of the owner of the firm and the unique tax code identifier. Thus, we use the platform to obtain these variables for out-of-state firms.

Matching owner names to cultural groups. We follow Bhagavatula et al. (2018) to match owner names to their Jatis (if the owners are of Hindu religion) or to their religion (in case the owners are not Hindu). Their procedure consists of using scraped data from Indian matrimonial websites that contain information on names, castes and religion. They train a sorting algorithm that uses names as inputs and gives a probability distribution across

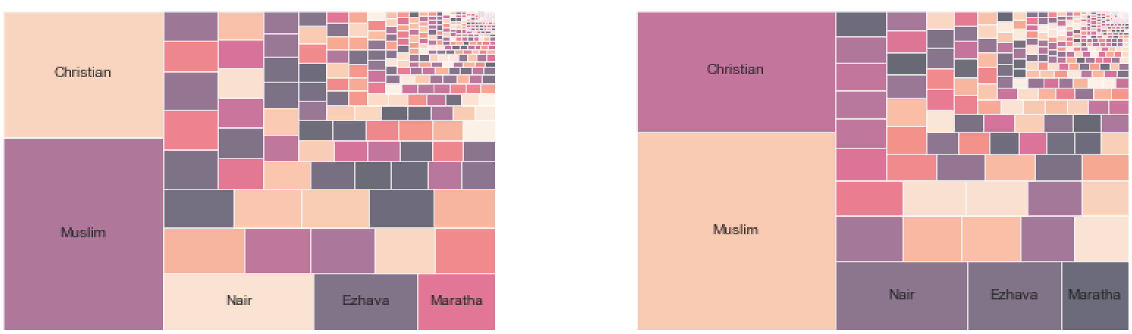
[8] `https://www.indiamart.com/`

cultural groups per name as outputs. We match these probability distributions to each owner name in our dataset. Notice that our notion of cultural group-belonging is probabilistic and not deterministic. This probabilistic approach is more relevant to our setup since, when firm owners trade with each other, they do not know each other's cultural group *ex ante*. Our sample finally consists of 452 cultural groups.

Merged dataset. For the analytical part we merge the three previous datasets. We end up with a sample that contains information from 22,295 unique firms, of which there are 10,559 sellers and 16,980 buyers. In total, the sample comprises approximately 560 thousand transactions or 97 billion rupees (around 1.4 billion US dollars). We drop any registered transaction in which the seller and the buyer is the same parent firm. Each firm is linked to a unique pincode. Finally, we assign a sector to each firm based on the HS codes of the goods sold. To provide a summary of the heterogeneity of cultural groups present in the firm-to-firm trade data, we show the distribution of probability-weighted sales and purchases across cultural groups in Figure 8.

Figure 8: Probability-weighted sales and purchases across cultural groups

(a) Sales (b) Purchases



Notes: Figure shows the decomposition of the probability-weighted sales and purchases across the 452 cultural groups in our dataset. The size of rectangles reflects the share of sales and purchases.

2.2.3 Construction of variables

Firm-to-firm trade variables. The firm-to-firm dataset provides information at the transaction level between any two registered firms. More specifically, we have information on (i) transaction-level unique identifiers, (ii) seller and buyer unique identifiers, (iii) the 6-digit HS description of the traded goods in each transaction, (iv) the total value of the transaction in rupees per type of good involved in each transaction and (v) the number of units sold of each good in each transaction.

For every seller/buyer pair we construct total sales, the total number of transactions, and unit values. For the total sales, we add up all the sales between each given pair of firms in our sample. We do the same with the total number of transactions. For obtaining the prices, we calculate the unit values. To do this, we first calculate the total amount sold and the total units sold of each good at the 6-digit HS level between each given pair of firms in our sample. Then, we divide the total amount sold by the number of units sold of each good.

Cultural proximity. Consider the set $\mathcal{X}$ of cultural groups, where $|\mathcal{X}| = X = 452$ in our final dataset. Since not all names are deterministically matched to a cultural group, each firm in our dataset has a discrete probability distribution over the set X of cultural groups. In particular, every firm ν has a probability distribution $\boldsymbol{\rho_\nu} = [\rho_\nu(1), \dots, \rho_\nu(X)]$, such that $\sum_{x=1}^{X} \rho_\nu(x) = 1$. In this part, we distinguish between the probability distribution over cultural groups of the seller and the probability distribution over cultural groups of the buyer. Define $\rho_\nu(x)$ as the probability of seller ν of belonging to cultural group x. Similarly, define $\rho_\omega(x)$ as the probability of buyer ω of belonging to cultural group x. Based on these two distributions we construct the following measure of cultural proximity: the Bhattacharyya (1943) coefficient.

The Bhattacharyya (1943) coefficient between seller ν and buyer ω measures the level of overlapping between two different probability distributions.[9] We define it as

$$BC(\nu, \omega) = \sum_{x=1}^{X} \sqrt{\rho_\nu(x)\, \rho_\omega(x)}.$$

Because $0 \leq \rho_\nu(x) \leq 1$ and $0 \leq \rho_\omega(x) \leq 1$, we have that $0 \leq BC(\nu, \omega) \leq 1$. On the one hand, $BC(\nu, \omega) = 0$ means the seller has a completely different probability distribution from that of the buyer. In our context, this means the seller and the buyer have no chance of belonging to the same cultural group or that their cultural proximity is the farthest. On the other hand, $BC(\nu, \omega) = 1$ means the seller has exactly the same probability distribution of the buyer. This implies that the seller has the same probability of belonging to a group of certain cultural groups than the buyer or that their cultural proximity is the closest possible.[10] In robustness checks, we use the Kullback and Leibler (1951) divergence measure

[9] Notice the Bhattacharyya coefficient is not the Bhattacharyya distance. The Bhattacharyya distance is defined as $BD(s, b) = -\log(BC(s, b))$. We prefer the Bhattacharyya coefficient because it is easier to interpret.

[10] For our purposes, it is important that the cultural proximity measure we use is symmetric. To see why, consider an example where, in our dataset, we have a transaction between a seller ν and a buyer ω, from which we obtain $BC(\nu, \omega)$. Further assume that in our dataset we record a second transaction in which the roles of the firms revert (i.e. the buyer becomes the seller and vice versa), so we calculate $BC(\omega, \nu)$. Regardless of the roles the firms take in this second transaction, we want their cultural proximity to remain constant, as

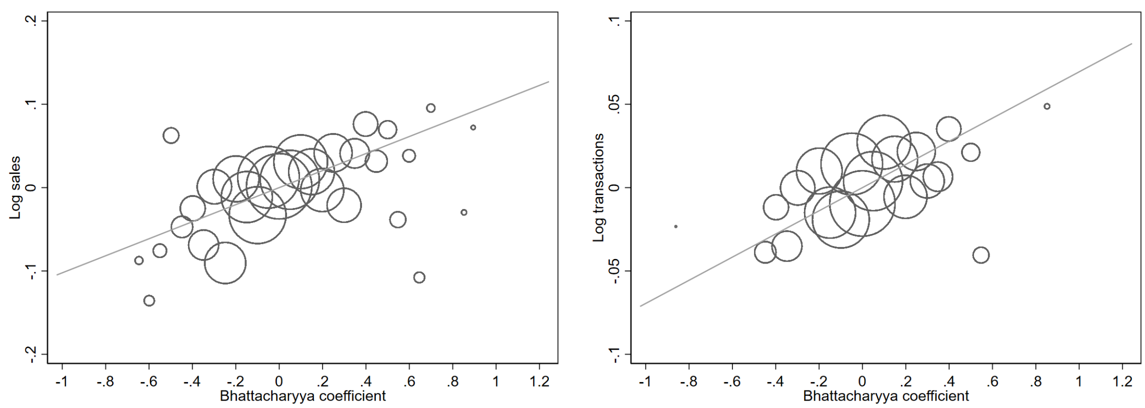
to measure cultural distance (Appendix B.3.1). All our results are qualitatively similar, and statistically significant when doing so.

2.3 Stylized facts

Fact 1: Cultural proximity fosters trade. We first discuss results related to the intensive margin of the firm-to-firm trade. Figure 9 shows the residualized scatterplots between the Bhattacharyya coefficient and two intensive margin measures: total sales between two firms and total transactions between two firms. The scatterplots show a higher Bhattacharyya coefficient (buyer and seller are probabilistically more alike in their cultural group) is related to a higher amount of sales and transactions.

Figure 9: Effect of cultural proximity on trade, intensive margin

(a) Sales (b) # Transactions



Notes: Results residualized of seller fixed effects, buyer fixed effects and log distance. Equally distanced bins formed over the X axis. Size of bubbles represents number of transactions in each bin. The higher the Bhattacharyya coefficient, the culturally closer two firms are.

the membership of cultural groups is fixed. This goal is achieved through the means of a symmetric proximity measure. Our example shows the Bhattacharyya coefficient complies with this symmetry requirement, as $BC(\nu, \omega) = BC(\omega, \nu)$.

Table 6: Effect of cultural proximity on trade, intensive and extensive margins

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable	Log Sales	Log Transactions	Log Sales	Log Transactions	Trade Indicator	Trade Indicator
BC	0.100***	0.066**	0.129***	0.076***	0.0009***	0.0010***
	(0.033)	(0.027)	(0.034)	(0.028)	(0.0001)	(0.0001)
Log dist.	-0.023	-0.065***			0.0001	
	(0.015)	(0.011)			(0.0000)	
Obs.	32,678	32,678	32,843	32,843	5,606,627	5,628,290
Adj. R2	0.415	0.359	0.410	0.356	0.617	0.0106
FE	Seller, buyer	Seller, buyer	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer	Seller, buyer, origin×dest.

Notes: Columns 1, 2, 3 and 4 show the results of estimating Equation (2.1). Columns 5 and 6 show the results of estimating Equation (2.2). ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Origin-destination fixed effect considers the district of the seller and the buyer. Standard errors two-way clustered at the seller and buyer level. Standard errors in parentheses. The higher the Bhattacharyya coefficient, the culturally closer two firms are. Number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia, Guimarães, and Zylkin 2019).

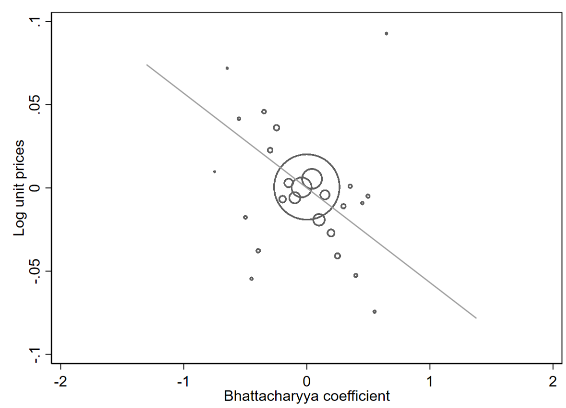
We now proceed to confirm the findings using a gravity equation. For transactions from firm ν to firm ω in our sample we estimate

$$\ln y(\nu,\omega) = \iota_\nu + \iota_\omega + \delta BC(\nu,\omega) + \eta \ln dist(\nu,\omega) + \varepsilon(\nu,\omega), \tag{2.1}$$

where $y(\nu,\omega)$ is either the total sales $n(\nu,\omega)$ or total transactions $t(\nu,\omega)$ from seller ν to buyer ω, $BC(\nu,\omega)$ is the Bhattacharyya coefficient, $dist(\nu,\omega)$ is the Euclidean distance between the pincodes in which the firms are located, ι_ν and ι_ω are seller and buyer fixed effects. Columns 1-4 of Table 6 present the results of the intensive margin estimation, which confirm the preliminary findings from Figure 9. Columns 1 and 2 show that, on average, there will be a higher amount of sales and transactions between a pair of firm when these firms are more alike in cultural terms. Columns 3 and 4 shows that these results remain strong after including origin-destination fixed effects, which account for geographic distance but also control for other features that might arise between a pair of locations such as different terrains, different languages, location-specific cultural ties, historical ties, etc.

Fact 2: Cultural proximity increases the likelihood of ever trading. Next, we estimate the extensive margin relationship. Given the size of our full dataset, the number of potential extensive margin links is computationally large. For tractability, we modify our sample. In the first place, we construct a sample with all possible combinations of in-state buyers and in-state sellers with cultural group information. Then, we proceed to drop all potential transactions that include unfeasible sectoral combinations. This means, we drop the combinations of firms that are involved in productive sectors that never recorded a transaction in the data. Finally, we drop all unfeasible transactions based on distance. This is to say, we drop the combinations of firms where the seller is further away than the maximum recorded distance for the in-state buyer or vice versa.

Figure 10: Effect of cultural proximity on prices



Notes: Results residualized of seller fixed effects and HS code fixed effects. Sectors defined according to 6-digit HS classification. Equally distanced bins formed over the X axis. Size of bubbles represents number of transactions in each bin. The higher the Bhattacharyya coefficient, the culturally closer two firms are.

With this sample, we construct a trade indicator variable $tr\left(\nu,\omega\right)$ which is equal to 1 if there is any kind of trade between firms ν and ω, and 0 otherwise. With this variable we estimate a gravity-type specification:

$$tr\left(\nu,\omega\right)=\iota+\iota_{\nu}+\iota_{\omega}+\delta BC\left(\nu,\omega\right)+\eta\ln dist\left(\nu,\omega\right)+\varepsilon\left(\nu,\omega,t\right). \tag{2.2}$$

Columns 5-6 of Table 6 present the extensive margin results. We find that the higher the Bhattacharyya coefficient, the more likely is that two given firms will trade.

Fact 3: Cultural proximity lowers prices. Figure 10 now uses buyer-seller-product groups and shows the residualized scatterplots between the similarity measures and the unit prices. We see the higher the Bhattacharyya coefficient between two firms involved in a transaction, the lower the price that will be charged. To confirm the results, we work with a seller-buyer-transaction-good version of our dataset and estimate

$$\ln p_{g}\left(\nu,\omega,t\right)=\iota_{\nu\times g}+\iota_{g\times t}+\iota_{\omega}+\delta BC\left(\nu,\omega\right)+\eta\ln dist\left(\nu,\omega\right)+\epsilon_{g}\left(\nu,\omega\right), \tag{2.3}$$

where $p_{g}\left(\nu,\omega,t\right)$ is the unit value of good g (at the 6-digit HS classification) sold by firm ν to firm ω in month t, $\iota_{\nu\times g}$ is a seller-good fixed effect and $\iota_{g\times t}$ is a good-month fixed effect.

We present the results in Table 7, which confirms the previous findings from the Figure: the closer the cultural proximity, the lower the unit value of the transactions.

Table 7: Effect of cultural proximity on prices

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable	Log Prices	Log Prices	Log Prices	Log Prices	Log Prices	Log Prices
BC	-0.069**	-0.069**	-0.066**	-0.045*	-0.040*	-0.039*
	(0.033)	(0.033)	(0.033)	(0.023)	(0.023)	(0.022)
Log dist.	0.023	0.023	0.028*			
	(0.016)	(0.016)	(0.017)			
Obs.	230,744	230,744	226,645	235,001	236,617	230,900
Adj. R2	0.932	0.932	0.935	0.933	0.925	0.936
FE	Seller×HS, buyer	Seller×HS, buyer, month	Seller×HS, buyer, month×HS	Seller×HS, buyer, origin×dest.	Seller×HS, buyer, month, origin×dest.	Seller×HS, buyer, month×HS, origin×dest.

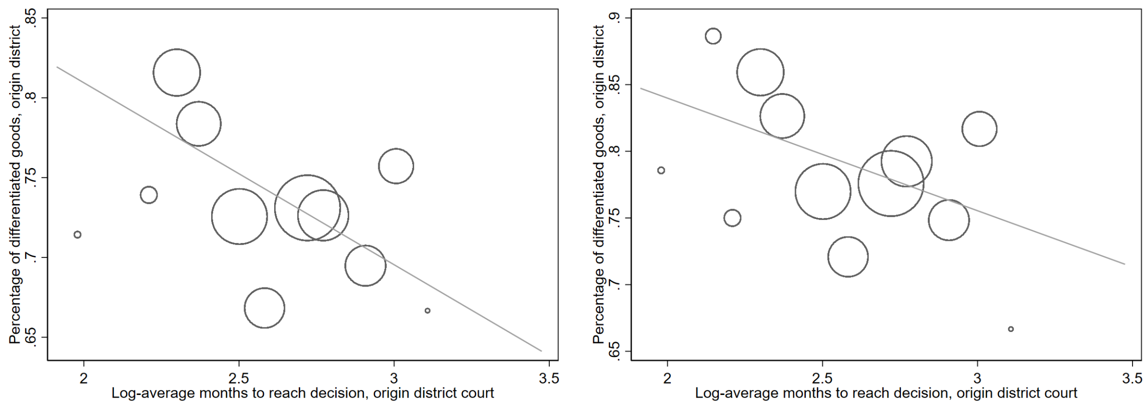
Notes: This table shows the results of estimating Equation (2.3). Good g is defined according to 6-digit HS classification. Prices trimmed by 4-digit HS code at 5 and 95 percent. ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Origin-destination fixed effect considers the district of the seller and the buyer. Standard errors are multi-way clustered at the seller, 4-digit HS and origin-destination level. Standard errors in parentheses. The higher the Bhattacharyya coefficient, the culturally closer two firms are. Number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia, Guimarães, and Zylkin 2019).

2.3.1 Differentiated goods and court quality

To better understand the underlying forces driving these empirical patterns, we explore the importance of contract enforcement, and cultural hierarchies. First, in this section, we show evidence that suggests that the effect we find of cultural proximity on trade is driven by differentiated goods, which often rely on either formal or informal contract enforcement (Nunn 2007). Then, we find that differentiated goods are more likely to be produced in and bought by firms that are located in districts with higher contract enforcement (as proxied by court delays). All in all, these analysis points that the stylized facts are likely driven by the desire of firms to reduce contracting frictions by trading with firms they trust. Here, cultural proximity arises as a proxy for knowing and trusting the other firm (Munshi 2014; Munshi 2019).

Figure 11: Differentiated goods and court quality by district

(a) Origin district (b) Destination district



Notes: Scatter plot at the district level. Equally distanced bins formed over the X axis. Size of bubbles represents number of observations in each bin. The larger the log-average number of months for cases to reach a decision, the worse the district's court. Differentiated goods according to the conservative classification of Rauch (1999). The log-average number of months for cases to reach a decision comes from Ash et al. (2021), where for each district court in the 2010-2018 dataset we take into account the average months in between a case's date of filing and date of decision.

In order to bring in information on the type of product, we first disaggregate our data at the seller-buyer-transaction-good level. Then, we classify the goods into differentiated goods and non-differentiated goods based on the classification developed by Rauch (1999).[11] We estimate the following specification:

$$\begin{aligned} \ln n_g(\nu,\omega,t) = \quad & \iota_{\nu\times g} + \iota_{g\times t} + \iota_\omega + \delta BC(\nu,\omega) + \xi\left(BC(\nu,\omega)\times \mathbb{I}_g^{diff}\right) \\ & +\eta \ln dist(\nu,\omega) + \epsilon_g(\nu,\omega), \end{aligned} \tag{2.4}$$

where $n_g(\nu,\omega,t)$ are the sales going from firm ν to firm ω of good g in month t and $\mathbb{I}_g^{diff}$ is an indicator for differentiated goods.[12] Table 8 presents the results for the sales. Our findings suggest that the baseline results of cultural proximity increasing trade are mostly driven by differentiated goods.

What could be the reason behind differentiated goods driving the cultural proximity results? At the international trade level Nunn (2007) suggests contract enforcement is related

[11]According to Rauch (1999) differentiated goods are the goods not traded in organized exchanges or not reference priced in commercial listings. Differentiated goods have specific characteristics that "differentiate" (i.e. specialized goods, branded goods) them from other more homogeneous types of goods. Because of their relative uniqueness in features, these goods are not as easily replaceable as non-differentiated goods and, as such, rely more on relationship-specific types of trade. This means sellers and buyers must face search frictions in order to match to a suitable trade partner and will likely not abandon the commercial matches they have already made.

[12]We use both the conservative and liberal classifications from Rauch (1999). The conservative classification minimizes the number of goods classified as non-differentiated and, thus, has the largest amount of differentiated goods. The liberal classification maximizes the amount of goods classified as differentiated and has the largest number of differentiated goods.

the production of relationship-specific goods. To analyze this, we construct a measure of court quality at the district level.[13] Using data from Ash et al. (2021) we calculate the log-average number of months for cases to reach a decision in each district court between 2010 and 2018. The larger the log-average number of months for cases to reach a decision, the worse quality this court has. Figure 11 shows that, in our dataset, districts with worse court quality sell and buy less differentiated goods, suggesting that differentiated products are more likely to be traded when contract enforcement is better.

Table 8: Effect of cultural proximity on trade by types of good, intensive margin

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable	Log Sales	Log Sales	Log Sales	Log Sales	Log Sales	Log Sales
BC	0.099***	0.018	0.039	0.069**	-0.019	0.013
	(0.031)	(0.050)	(0.040)	(0.027)	(0.048)	(0.038)
$BC \times \mathbb{I}_g^{diff,con}$		0.122**			0.139**	
		(0.058)			(0.059)	
$BC \times \mathbb{I}_g^{diff,lib}$			0.097**			0.095**
			(0.047)			(0.047)
Obs.	174,352	174,352	174,352	177,584	177,584	177,584
Adj. R2	0.852	0.852	0.852	0.853	0.853	0.853
FE	Seller×HS, buyer, month×HS	Seller×HS, buyer, month×HS	Seller×HS, buyer, month×HS	Seller×HS, buyer, month×HS, origin×dest.	Seller×HS, buyer, month×HS, origin×dest.	Seller×HS, buyer, month×HS, origin×dest.

Notes: This table shows the results of estimating Equation (2.4). ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Good g is defined according to 6-digit HS classification. Sales trimmed by 4-digit HS code at 5 and 95 percent. Origin-destination fixed effect considers the district of the seller and the buyer. Standard errors two-way clustered at the seller and 4-digit HS level. Standard errors in parentheses. The higher the Bhattacharyya coefficient, the culturally closer two firms are. Number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia, Guimarães, and Zylkin 2019). $\mathbb{I}_g^{diff,con}$ indicates the good g is a differentiated one according to the conservative classification of Rauch (1999). $\mathbb{I}_g^{diff,lib}$ indicates the good g is a differentiated one according to the liberal classification of Rauch (1999).

Following an argument similar to Munshi (2019) and Nunn (2007), we interpret these findings as evidence that cultural proximity relates to contract enforcement and trust. Differentiated goods do not trade in exchanges and are not homogeneous, but are branded and specific to certain producing firms. In a country with market imperfections as India, firms can easily renege on their commitments. For buyers this could be not much of a hassle when it comes to homogeneous goods, as their suppliers are easily interchangeable. For sellers, this could be just a small problem as they can easily find other buyers. However, problems can arise if firms renege on their commitments related to differentiated goods. Suppliers

[13] See Ash et al. (2021), Boehm and Oberfield (2020), and Rao (2019) for references that analyze the effects of court quality in India.

and buyers in differentiated good markets are not easily replaceable. As a result suppliers of differentiated goods will only sell to buyers that they know and trust, while buyers of differentiated goods will do the same when choosing sellers.[14] Therefore, in these cases, trade will increase when firms trust and know each other, that is, when firms are culturally close.

2.3.2 Hierarchies

To investigate the importance of vertical hierarchies and discrimination across cultural groups, we study whether there are asymmetric effects in transactions in which one firm is placed higher than the other based on the Varna-based hierarchy. This is one way of testing for preference-based discrimination across the social hierarchy. We generate indicators based on which is the Varna or religion for which a firm has the highest probability of belonging to.[15] We do not find evidence that hierarchies (and preference-based discrimination) across social groups matter for our social proximity results.

We make use of two different indicators: $\mathbb{I}_{\nu_H \omega_L}$ and $\mathbb{I}_{\nu_L \omega_H}$. The first one indicates that the seller belongs to a higher hierarchy than the buyer. The second one indicates the buyer is placed below the seller in the social hierarchy. We include these two indicators by interacting them with our measure of cultural proximity. Table 9 presents the results for the intensive and extensive margins. The baseline category is that both firms belong to the same hierarchy. First place, we find the baseline coefficient is very similar to those of Table 6. Second, we find there is no additional effect of cultural proximity when firms are placed differently in the hierarchy. We conclude that strong asymmetric effects caused by vertical discrimination across cultural groups are unlikely. The effect of cultural proximity is similar, whether or not the firms trading belong to the same or different hierarchies.

[14] We can relate our result to that of Rauch (1999), who mentions that search frictions (i.e. having to look for a trustworthy supplier) are more important to the trade of differentiated goods than to the trade of non-differentiated goods.

[15] While the Varna-based hierarchy only relates to the Hindu religion, we also place other religions in this hierarchy based on their income levels. We do this to prevent losing a large share of the sample in our estimations.

Table 9: Effect of cultural proximity on trade by vertical hierarchies, intensive and extensive margins

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable	Log Sales	Log Transactions	Log Sales	Log Transactions	Trade Indicator	Trade Indicator
BC	0.099***	0.068**	0.129***	0.079***	0.0010***	0.0010***
	(0.034)	(0.028)	(0.035)	(0.029)	(0.0001)	(0.0001)
$BC \times \mathbb{I}_{\nu_H \omega_L}$	0.023	0.097	0.008	0.072	-0.0002	-0.0003
	(0.113)	(0.091)	(0.116)	(0.092)	(0.0003)	(0.0003)
$BC \times \mathbb{I}_{\nu_L \omega_H}$	0.045	-0.076	-0.027	-0.123	-0.0002	-0.0004
	(0.128)	(0.102)	(0.129)	(0.103)	(0.0002)	(0.0002)
Obs.	30,997	30,997	31,119	31,119	5,456,512	5,477,548
Adj. R2	0.418	0.360	0.412	0.357	0.614	0.0107
FE	Seller, buyer	Seller, buyer	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer	Seller, buyer, origin×dest.

Notes: Columns 1, 2, 3 and 4 show the results of estimating a modified version of Equation (2.1). Columns 5 and 6 show the results of estimating a modified version of Equation (2.2). ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Origin-destination fixed effect considers the district of the seller and the buyer. Standard errors two-way clustered at the seller and buyer level. Standard errors in parentheses. The higher the Bhattacharyya coefficient, the culturally closer two firms are. Number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia, Guimarães, and Zylkin 2019). The subindex that accompanies ν denotes the hierarchical position of the seller, while the subindex that accompanies ω denotes the hierarchical position of the buyer. H denotes a higher position and L denotes a lower position. The baseline category is when both firms have the same hierarchical position.

2.3.3 Additional specifications

We examine alternative specifications and heterogeneity in responses that shed light on various other channels in Appendix B.3.

Alternative cultural proximity measure. As an alternative to the Bhattacharyya coefficient, we perform estimation exercises using a symmetric version of the Kullback and Leibler (1951) divergence. Tables B3 and B4 show our baseline findings are robust to this alternative cultural proximity measure.

Language. We test whether the results we find are driven by language similarity. To do so, we follow the two linguistic distance measures from **kone2018internal**. Table B5 shows that language does not affect the cultural proximity results already established.

Goods specialization. Cultural groups in India are, in many cases, defined by the production of specific goods (Munshi 2019).[16] Therefore, we analyze if the reason behind the cultural proximity results is cultural groups specializing in the production of certain goods

[16] We can also understand this as certain cultural groups specializing in certain occupations.

and, given this, forming special bonds with their specific set of buyers. In Table B6 we do not find evidence of good specialization driving the results. This means that cultural proximity matters for all types of goods: for those in which a cultural group specializes and for those in which a cultural group does not specialize too.

Number of varieties sold and bought. We analyze whether firms that the social proximity results prevail for firms that sell and buy more varieties goods. To measure this, we count how many varieties of inputs a firm buys or how many varieties of goods a firm sells. In Table B7 we find the more varieties a firm sells or buys, the more the intensity of trade is affected by social proximity. Our interpretation of these findings is that the more varieties a firm sells or buys, the more contracting frictions it faces, caused by having to negotiate with either more suppliers or more clients. These firms, in order to minimize their load of contracting frictions, will rely more on trading with counterparts in which they trust (i.e. firms that are culturally close).

2.3.4 Discussion of stylized facts

The stylized facts show that a higher cultural proximity between a pair of firms favors trade in both the intensive and extensive margins, as well as lowers the price of the goods they trade. We discuss the possible mechanisms that may give rise to these findings.

Contracting frictions. In Section 2.3.1 we argue that contracting frictions could be the reason that drives the cultural proximity results. India is a country that suffers from severe lack of contract enforcement. *A priori*, a buyer may not know if the seller will deliver the goods under the agreed conditions (delivery, quality, etc.). Likewise, *a priori*, the seller may not know if the buyer will pay under the agreed conditions. This means buyers and sellers incur contracting frictions to find suitable trading schemes or partners (Boehm and Oberfield 2020). Quantity-wise and matching-wise, this lowers trade as firms must pay a matching cost. Price-wise, this increases prices as the matching cost is passed down by the sellers to the aforementioned prices.

In this case, cultural proximity can work as a proxy for information and trust: culturally close firms may know and/or trust each other, and/or informally enforce contracts with social and reputational pressures. The higher the cultural proximity, the lower the contracting frictions. Therefore, there would be more trade and lower prices, which is consistent with our previous findings. In Section 2.4 we present a simple theoretical framework in which cultural proximity affects contracting frictions and affects trade and prices. While our model is agnostic about why cultural proximity bridges the wedges in prices, the above discussion suggests that if contracting frictions drive initial trade barriers, then cultural proximity may reduce such frictions.

Preference-based mechanisms. We argue the results are unlikely to emerge from buyers having an inherent preference for buying from sellers culturally close to them. We could model this preference as a demand shifter that is active for those sellers that are close in cultural terms. While this would certainly increase the quantity traded, it would increase the price of traded goods, a result that is not consistent with our previous findings.

The stylized facts can arise from having sellers that show a preference for selling to buyers that are culturally close. It would imply the introduction of a supply shifter that is active for those buyers that are culturally close to the seller. Yet, this channel is unlikely, as in the presence of profit maximizing firms, such firms may be competed out of the market. Discrimination from high-caste cultural groups against low-caste cultural groups may again reduce trade. Yet, in Section 2.3.2 we find this to be an unlikely driver of our empirical patterns. That is, we find there is no additional effect of cultural proximity when firms are placed differently in the hierarchy. As such, we detect no asymmetric effects caused by vertical discrimination across cultural groups.

2.4 Model

In this section we describe the model environment and define the equilibrium of the model, and Appendix B.4 contains further details.

2.4.1 Environment

Following Bernard et al. (2022), we build a quantitative firm-level production network model with heterogeneous firms and endogenous network formation. We modify the original setting to not only make firms heterogeneous in productivity, but also in their cultural endowments. We use these cultural endowments to construct a measure of cultural proximity between firms, which in turn influences trade costs and matching costs.

Firms. There is a continuum of firms in the economy that operate under monopolistic competition and produce differentiated goods indexed by ω. We consider a roundabout production economy, so each firm produces by hiring labor from a representative household and by purchasing intermediate inputs from all the other firms in the economy.

Demand for firms comes from two different sources. First, as mentioned, the output of each firm is demanded by other firms as intermediate inputs. Second, the output of each firm is demanded by a representative household as consumption goods. Firms charge the same price for its differentiated output to both households and the rest of firms.

Each firm has a technology

$$y\left(\omega\right)=\kappa_{\alpha}z\left(\omega\right)l\left(\omega\right)^{\alpha}m\left(\omega\right)^{1-\alpha}, \tag{2.5}$$

where $y\left(\omega\right)$ is output, $\kappa_{\alpha}\equiv\frac{1}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}$ is a normalization constant, $z\left(\omega\right)$ is firm-level productivity, $l\left(\omega\right)$ is labor, and $m\left(\omega\right)$ are intermediate inputs from other firms. In turn, the intermediate inputs are defined as a CES composite so

$$m\left(\omega\right)=\left(\int_{\nu\in\Omega(\omega)}m\left(\nu,\omega\right)^{\frac{\sigma-1}{\sigma}}d\nu\right)^{\frac{\sigma}{\sigma-1}},$$

where $m\left(\nu,\omega\right)$ is quantity of inputs from seller ν to buyer ω, $\sigma>1$ is the elasticity of substitution across intermediates, and $\Omega\left(\omega\right)$ is the endogenous set of suppliers of buyer ω. By cost minimization we get

$$c\left(\omega\right)=\frac{P\left(\omega\right)^{1-\alpha}}{z\left(\omega\right)}, \tag{2.6}$$

where $P\left(\omega\right)\equiv\left(\int_{\nu\in\Omega(\omega)}p\left(\nu,\omega\right)^{1-\sigma}d\nu\right)^{\frac{1}{1-\sigma}}$ is a CES price index across prices of intermediates, and labor is the numeraire good, so $w=1$. Profit maximization subject to demand generates constant markup pricing such that

$$p\left(\nu,\omega\right)=\mu c\left(\nu\right)d\left(\nu,\omega\right),\ d\left(\nu,\omega\right)\geq1, \tag{2.7}$$

where $d\left(\nu,\omega\right)$ is a pricing wedge that increases the price that seller ν charges to buyer ω, and $\mu\equiv\frac{\sigma}{\sigma-1}$ is the markup. We will define this wedge in the following paragraphs. We now derive the demand for intermediates, so

$$n\left(\nu,\omega\right)=p\left(\nu,\omega\right)^{1-\sigma}P\left(\omega\right)^{\sigma-1}N\left(\omega\right), \tag{2.8}$$

where $N\left(\omega\right)=\int_{\nu\in\Omega(\omega)}n\left(\nu,\omega\right)d\nu$ is the total intermediate purchases by buyer ω and $n\left(\nu,\omega\right)\equiv p\left(\nu,\omega\right)m\left(\nu,\omega\right)$ is the value of purchases from seller ν to buyer ω. From Equation (2.8) we can obtain the gravity equation as

$$\ln\left(n\left(\nu,\omega\right)\right)=\iota_{\nu}+\iota_{\omega}+\left(1-\sigma\right)\log\left(d\left(\nu,\omega\right)\right), \tag{2.9}$$

where ι_{ν} and ι_{ω} are seller and buyer fixed effects. This gravity equation relates directly to

Equation (2.1). Lastly, we assume the wedge is a function of different trade costs, including cultural proximity between firm owners due to ethnicity. Thus, we have

$$d(\nu, \omega) = \exp\left(\beta_1 dist(\nu, \omega) + \beta_2 BC(\nu, \omega)\right), \tag{2.10}$$

where the parameters β_1 and β_2 are trade cost semi-elasticities. The wedge will be larger the longer the geographic distance and the lower the cultural proximity. From Equation (2.7) we have that the higher the cultural proximity, the lower the prices, which relates to stylized Fact 3. Likewise, from Equation (2.9) we have that the higher the cultural proximity, the higher the intermediate sales, which relates to stylized Fact 1.

Households. There is a representative household that demands goods from firms and inelastically supplies labor to them. To simplify, the representative household exhibits the same elasticity of substitution across goods σ as from firms. So, the representative household solves

$$\max_{\{y(\omega)\}} \left(\int_{\omega\in\Omega} y(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right)^{\frac{\sigma}{\sigma-1}}, \; s.t. \int_{\omega\in\Omega} P(\omega)\, y(\omega)\, d\omega \leq Y,$$

where $P(\omega)$ is the price the household pays for good sold by ω, Ω is the set of firms in the economy, and Y is total income. This generates the demand for good ω

$$x(\omega) = P(\omega)^{1-\sigma} P^{\sigma-1} Y, \tag{2.11}$$

where $x(\omega) \equiv P(\omega)\, y(\omega)$ is the value of purchases from ω, and $P \equiv \left(\int_{\omega\in\Omega} P(\omega)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ is a CES price index.

2.4.2 Equilibrium given production network

Here, we lay out the equilibrium conditions conditional on the structure of the network. Conditional on the formation of the network, firms only differ in productivity z, so we now identify each firm according to its productivity. Based on the price index of all of the goods acquired by firm z', we get

$$P\left(z'\right)^{1-\sigma} = \mu^{1-\sigma} \int P(z)^{(1-\alpha)(1-\sigma)} z^{\sigma-1} d\left(z, z'\right)^{1-\sigma} l\left(z, z'\right) dG(z), \tag{2.12}$$

where $l\left(z, z'\right)$ is the share of sellers of productivity z that sell to buyers with productivity z', also called the *link function*. Now, total sales of firm z is the sum of sales to household plus intermediates, so

$$S\left(z\right)=\frac{\left[\mu^{1-\sigma}P\left(z\right)^{(1-\alpha)(1-\sigma)}z^{\sigma-1}\right]\times}{\left[\frac{Y}{P^{1-\sigma}}D\left(z\right)^{1-\sigma}+\left(\frac{1-\alpha}{\mu}\right)\left(\int\left[d\left(z,z^{'}\right)^{1-\sigma}P\left(z^{'}\right)^{\sigma-1}S\left(z^{'}\right)\right]l\left(z,z^{'}\right)dG\left(z^{'}\right)\right)\right]},\tag{2.13}$$

where $D\left(z\right)=\int_{\omega\in\Omega(\nu)}d\left(\nu,\omega\right)d\omega=\int d\left(z,z^{'}\right)l\left(z,z^{'}\right)dG\left(z^{'}\right)$ is the aggregated wedge for firm of productivity z.

2.4.3 Endogenous network

We endogeneize the formation of the production network by laying out the maximization problem of firms and how cultural proximity influences it. In particular, we allow for the cost of sellers and buyers matching to depend on their cultural proximity, which we can then estimate from the data. Before the formation of the network, firms are characterized by the tuple $\lambda=(z,\boldsymbol{\rho})$, where z is productivity, and $\boldsymbol{\rho}$ is the vector of probabilities of firm λ belonging to each cultural group. We can then construct a measure of cultural proximity according to the Bhattacharyya coefficient, such that

$$BC\left(z,z^{'}\right)=\sqrt{\sum_{x}\rho_{z}\left(x\right)\rho_{z^{'}}\left(x\right)}.$$

Now we describe how firms match. A seller z trades with a buyer $z^{'}$ only if it is profitable for the seller to do. To trade, the seller incurs in a pairwise matching cost $F\left(z,z^{'}\right)$.[17] Then, the share of seller-buyer pairs $\left(z,z^{'}\right)$ is

$$l\left(z,z^{'}\right)=\int\mathbb{I}\left[\ln\left(\pi\left(z,z^{'}\right)\right)-\ln\left(F\left(z,z^{'}\right)\right)-\ln\left(\epsilon\left(z,z^{'}\right)\right)>0\right]dH\left(\epsilon\left(z,z^{'}\right)\right),\tag{2.14}$$

where $\pi\left(z,z^{'}\right)$ are the profits for seller z of selling to buyer $z^{'}$ and ϵ is an i.i.d. log-normal noise variable with mean 0 and standard deviation $\sigma_{\ln(\epsilon)}$. Intuitively, the link function can be understood as the probability a seller z will match to a seller $z^{'}$. We define the pairwise matching cost to be related to the cultural distance. Then

$$F\left(z,z^{'}\right)=\kappa+\exp\left(\gamma BC\left(z,z^{'}\right)\right),\tag{2.15}$$

where γ measures the sensitivity of the pairwise matching cost to the cultural distance

[17] We assume that the matching cost is paid by the seller. For a further discussion on the importance of whether the seller or the buyer pays the fixed cost, see Huneeus (2018).

and κ is a scaling constant. From Equations (2.14) and (2.15), we see that the higher the cultural proximity, the lower the matching cost and the larger the probability of matching. This relates to stylized Fact 2.

2.5 Estimation and calibration

Here we explain how we estimate the key parameters of the model on cultural endowments, (intensive) trade costs, and seller matching costs. We also describe how calibrate the remaining parameters of the model.

Cultural endowments $\boldsymbol{\rho}$. For the cultural endowments, we assume each firm ν has a probability vector $\boldsymbol{\rho_\nu} = [\rho_\nu(1), \ldots, \rho_\nu(452)]$ of belonging to each of the 452 cultural groups we observe in the data. We further assume the elements of $\boldsymbol{\rho_\nu}$ are randomly drawn from a Dirichlet distribution, such that $\rho_\nu(1), \ldots, \rho_\nu(452) \sim \mathcal{D}(\alpha_1, \ldots, \alpha_{452})$, where $\alpha_1, \ldots, \alpha_{452} > 0$ are concentration parameters.[18] The probability density function for the Dirichlet distribution is

$$\rho_\nu(1), \ldots, \rho_\nu(452) \sim \mathcal{D}(\alpha_1, \ldots, \alpha_{452}) = \frac{\Gamma\left(\sum_{x=1}^{452} \alpha_x\right)}{\prod_{x=1}^{452} \Gamma(\alpha_x)} \prod_{k=1}^{452} \rho_\nu(x)^{\alpha_x - 1},$$

such that $\rho_\nu(x) \in [0,1]$, $\quad \sum_{x=1}^{452} \rho_\nu(x) = 1$, where $\Gamma(.)$ is the gamma function and $\frac{\Gamma\left(\sum_{x=1}^{452} \alpha_x\right)}{\prod_{x=1}^{452} \Gamma(\alpha_x)}$ is a normalization constant. To ensure the theoretical Dirichlet distribution produces draws that are similar to the probabilities we see in the data, we estimate the vector $\boldsymbol{\alpha} = \left[\begin{array}{ccc} \alpha_1, & \ldots, & \alpha_{452} \end{array}\right]$ parameters by maximum likelihood.[19] Let $\boldsymbol{\varrho} = \{\boldsymbol{\rho_1}, \ldots, \boldsymbol{\rho_N}\}$, where $\mathcal{N}$ is the total number of firms. Then, the log-likelihood function is

$$\ln pr(\boldsymbol{\varrho}|\boldsymbol{\alpha}) = \mathcal{N} \ln \Gamma\left(\sum_{x=1}^{452} \alpha_x\right) - \mathcal{N} \sum_{x=1}^{452} \ln \Gamma(\alpha_x) + \mathcal{N} \sum_{x=1}^{452} (\alpha_x - 1)\left(\frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \ln \rho_\nu(x)\right). \tag{2.16}$$

Trade costs d. From Equation (2.10) we need an estimate for for $\{\beta_1, \beta_2\}$. We obtain estimates for these two parameters by linking the theoretical gravity equation (2.9) to the empirical gravity equation results (Column 1 from Table 9). Thus, we obtain $\{\beta_1, \beta_2\} = \{0, -0.03\}$.[20]

[18] For a given x, the higher this parameter, the more disperse the realizations of $\rho_\nu(x)$ are across firms ν.

[19] For this, we use the Matlab toolboxes `fastfit` and `lightspeed` by Tom Minka. We present the estimated parameters in Figure B1 in Appendix A.1.

[20] Even though the wedge also appears in the price equation 2.7 of the model, we do not estimate this equation to identify β_1 and β_2. The reason is that the price equation is not an equilibrium equation, while the gravity equation is. Also, for our simulations we add a constant to the trade cost, such that the minimum

Matching cost F. From Equation (2.15), we need an estimate for γ. We do this in two steps. First, using the extensive margin sample we run the following estimation

$$\ln\left[n\left(z,z^{'}\right)\right]=\iota_z+\iota_{z^{'}}+\delta BC\left(z,z^{'}\right)+\gamma\ln\left(dist\left(z,z^{'}\right)\right)+\varepsilon\left(z,z^{'}\right),\tag{2.17}$$

where we apply the inverse hyperbolic sine transformation to the dependent variable, so as to not lose the cases in which there is zero trade. With this we recover

$$\ln\widehat{[n\left(z,z'\right)]}=\widehat{\iota}_z+\widehat{\iota}_{z'}+\widehat{\delta}BC\left(z,z^{'}\right)+\widehat{\eta}\ln\left(dist\left(z,z^{'}\right)\right),$$

where the hats denote estimated parameters and $\ln\widehat{[n\left(z,z'\right)]}$ are the predicted sales. This variable predicts what would be the sales for a pair of seller and buyer even in the case they did not actually trade in the data. Second, we combine and rearrange Equations (2.14) and (2.15), such that

$$l\left(z,z^{'}\right)=\int 1\left[\ln\left(\epsilon\left(z,z^{'}\right)\right)<\ln\widehat{[n\left(z,z'\right)]}-\ln\left(\sigma\right)-\gamma BC\left(z,z^{'}\right)\right]dH\left(\epsilon\left(z,z^{'}\right)\right),\tag{2.18}$$

where we use the fact that $\pi\left(z,z^{'}\right)=\frac{n\left(z,z^{'}\right)}{\sigma}$ and replace $\ln\left[n\left(z,z^{'}\right)\right]$ by its estimated counterpart $\ln\widehat{[n\left(z,z'\right)]}$.[21] We estimate this last equation with a probit regression (assuming $\epsilon\left(z,z^{'}\right)$is log-normally distributed). We find that $\gamma=-0.13$.[22]

Calibrated parameters and SMM. We calibrate the labor cost share $\alpha=0.52$, the value reported for India for 2019 from the Penn World Tables (Feenstra, Inklaar, and Timmer 2015). This value also considers the informal sector, which plays a large role in India. For the markup we use $\mu=1.34$, which is the median markup across all Indian sectors reported by (De Loecker et al. 2016). This markup implies an elasticity of substitution across suppliers $\sigma=3.94$. Following Bernard et al. (2022) we normalize the total number of workers $L=1$, take the nominal wage as the numeraire so $w=1$, and set the total number of firms $\mathcal{N}=400$.

For the log-productivity distribution, we assume a mean $\mu_{\ln(z)}=0$. The remaining parameters are (i) the standard deviation of the log-productivity distribution $\sigma_{\ln(z)}$ and (ii) the mean $\mu_{\ln(\epsilon)}$, (iii) the standard deviation $\sigma_{\ln(\epsilon)}$ of the link function noise distribution and (iv) the scaling constant for the pairwise matching cost κ. We estimate these four

trade cost is equal to 1. Therefore, in our simulations we have $d\left(\nu,\omega\right)=\exp\left(-\beta_2+\beta_2 BC\left(\nu,\omega\right)\right)$.

[21] For these estimations we ignore the scaling constant κ that appears in Equation (2.15).

[22] We present the results of the estimation in Table B1 in Appendix A.1. Also, for our simulations we add a constant to the matching cost, such that the minimum matching cost is equal to κ. Therefore, in our simulations we have $F\left(z,z^{'}\right)=\kappa+\exp\left(-\gamma+\gamma BC\left(z,z^{'}\right)\right)$.

parameters so as to match targeted moments from the data, using a simulated method of moments (SMM). We explain this procedure below.

Targeted and untargeted moments. Since the link function noise distribution affects how firms match between them, to identify the parameters related to this distribution we must target moments that are related to the extensive margin.

First, we choose to target the mean of the log-normalized number of buyers $\ln\left(\frac{\mathcal{N}_b(\nu)}{\mathcal{N}}\right)$, where $\mathcal{N}_b(\nu)$ is the number of buyers a seller ν has; and the mean of the log-normalized number of sellers $\ln\left(\frac{\mathcal{N}_s(\omega)}{\mathcal{N}}\right)$, where $\mathcal{N}_s(\omega)$ is the number of sellers a buyer ω has. Because these two moments are related to magnitude of the matching, they should inform us about the mean of the link function noise distribution $\mu_{\ln(\epsilon)}$ and the scaling constant for the pairwise matching cost κ.

Second, this being mostly a seller-oriented model, to identify the standard deviation of the link function noise distribution $\sigma_{\ln(\epsilon)}$ we target the variance of the log-normalized number of buyers $\ln\left(\frac{\mathcal{N}_b(\nu)}{\mathcal{N}}\right)$. Lastly, to identify the standard deviation of the log-productivity distribution, we must choose a moment that is related to the variance of the intensive margin. Thus, we target the variance of the log-normalized intermediate sales $\ln\left(\frac{\widetilde{N}(\nu)}{\mathcal{N}_b(\nu)}\right)$, where $\widetilde{N}(\nu)$ is the total intermediate sales a seller ν makes.

The first untargeted moment we consider is the variance of the log-normalized number of sellers $\ln\left(\frac{\mathcal{N}_s(\omega)}{\mathcal{N}}\right)$. The second untargeted moment we examine is the variance of the log-normalized intermediate purchases $\ln\left(\frac{N(\omega)}{\mathcal{N}_s(\omega)}\right)$. The exact definition of the targeted and untargeted moments, as well as the construction of their empirical counterparts, appears in Appendix B.2.

Goodness of fit. After our matching procedure, we find the parameters $\sigma_{\ln(z)} = 0.88$, $\mu_{\ln(\epsilon)} = 64.30$, $\sigma_{\ln(\epsilon)} = 10.85$ and $\kappa = 14.80$. Table B2 in Appendix B.2 shows how the model-based moments fare against their empirical counterparts. When it comes to the targeted moments, the model can very closely replicate the empirical ones. For the untargeted moments, the model gets reasonably close to the data.

2.6 Counterfactuals

We now present the results of various counterfactual exercises. First, we evaluate the effects of social mixing/inclusion and isolation policies, such that we change the cultural proximity between firms (in our model terms, changing $BC\left(z, z'\right)$). Second, we study the effects of a policy that reduces contracting frictions, such that firms rely less on cultural proximity when trading (in terms of our model, shrinking parameters β_2 and γ).

To evaluate each scenario, we measure what happens to various model-based statistics. Welfare is measured by real wage, $\mathcal{W} = \frac{w}{P}$. To quantify the impact on aggregate productivity, we consider a sales-weighted average productivity measure such that $\mathcal{Z} = \left(\sum_{\nu=1}^{N} \phi_\nu z_\nu^{\sigma-1}\right)^{\frac{1}{\sigma-1}}$, where ϕ_ν represents the proportion of the sales of firm ν over the total sales of the economy. To analyze the impact on the total economic activity, we measure total sales $\mathcal{S} = \sum_{\nu=1}^{N} S_\nu$, where S_ν are the total sales of firm ν. Additionally, we consider the average normalized intermediate sales $mean\left[\ln\left(\widetilde{N}\left(\nu\right)/\mathcal{N}_b\left(\nu\right)\right)\right]$, where $\widetilde{N}\left(\nu\right)$ are the total intermediate sales of seller ν, and the average normalized intermediate purchases $mean\left[\ln\left(N\left(\omega\right)/\mathcal{N}_s\left(\omega\right)\right)\right]$. For the prices, we compare the changes in the aggregate price index P. Finally, to study how matching between firms is affected, we present the results for the average normalized number of buyers, $mean\left[\ln\left(\frac{\mathcal{N}_b(\nu)}{\mathcal{N}}\right)\right]$, and the average normalized number of sellers, $mean\left[\ln\left(\frac{\mathcal{N}_s(\omega)}{\mathcal{N}}\right)\right]$.[23]

2.6.1 Social inclusion and social mixing policies

We analyze the effects of social inclusion or social mixing policies.[24] For instance, affirmative actions programs may help incentivize students from different cultural groups to attend the same educative institutions. If these students then go on to become owners of the firms in the future, such policies may increase cultural proximity between these firms, despite the fact the owners originally belonged to different cultural groups. To analyze the maximum potential of this policy within our theoretical framework, we propose case Counterfactual 1 (CF1) in which all the firms belong to the same cultural group. This is, we go from the baseline to $BC\left(z, z'\right) = 1$ for all z, z', which makes the firms to become the closest possible in cultural terms. In this scenario, there are no contracting frictions, as firms know and/or trust each other, and so they pay the minimum trade and matching costs.

Table 10 shows how the model statistics change in each counterfactual with respect to the baseline. In case CF1, we have that firms become the closest in cultural terms, so trade costs and matching costs go to their minimum possible. Aligned with our empirical facts, with lower trade costs, total sales increase by 2.76 percent, while the average intermediate sales and purchases go up by 1.52 percent and 1.15 percent, respectively. With the lower matching costs the average number of buyers grows by 1.07 percent, and the average number of sellers goes up by 1.00 percent. Also, because there are lower trade and matching costs, aggregate

[23]In contrast to the previous sections, in this part we define the aggregate measures discretely. This is due to the simulations having a discrete number of firms, rather than a continuum.

[24]See Munshi (2019) for a brief discussion of policies put forward in India to diminish the effects of castes through education.

prices fall by 1.73 percent. With this, welfare increases by 1.76 percent. Besides welfare, another aggregate measure we analyze is average productivity, which falls by 0.13 percent. Yet, average productivity masks substantial compositional changes, as these results depend on whether the less productive firms are selling more or less with respect to the baseline case. We show in Table 11 that, in case CF1, when trade and matching costs decrease, the less productive firms match more and sell more, which increases their weight in the aggregate and lowers average productivity.

Table 10: Effect of cultural proximity on aggregate outcomes (counterfactual scenarios)

	CF1: Social inclusion/mixing	CF2: Social isolation	CF3: Reducing contracting frictions
Welfare	1.76	-1.45	0.87
Ave. productivity	-0.13	0.10	-0.06
Total sales	2.76	-2.23	1.37
Ave. normalized intermediate sales	1.52	-1.20	0.76
Ave. normalized intermediate purchases	1.15	-0.94	0.57
Ave. normalized number of buyers	1.07	-0.87	0.53
Ave. normalized number of sellers	1.00	-0.82	0.50
Agg. price index	-1.73	1.47	-0.87

Notes: We present the percentage gains or losses with respect to the baseline scenario. CF1 is a case where all the firms belong to the same cultural group. This is, we go from the baseline to $BC\left(z, z^{'}\right) = 1$ for all $z, z^{'}$, which makes the firms to become the closest possible in cultural terms. In this scenario, there are no contracting frictions, as firms know and/or trust each other, and so they pay the minimum trade and matching costs. CF2 is a case where each firm belongs to its own cultural group. Thus, we have a case where $BC\left(z, z^{'}\right) = 0$ for all $z, z^{'}$ and $z \neq z^{'}$, which makes the firms the furthest possible in cultural terms. Under this scenario, firms incur the maximum contracting frictions, for which they pay the maximum trade cost and the maximum matching cost. CF3 is a scenario where trade and matching costs become less sensitive to cultural proximity. In this case parameters β_2 and γ shrink by 50 percent.

Table 11: Change in sales by productivity quartiles

	CF1: Social inclusion/mixing	CF2: Social isolation	CF3: Reducing contracting frictions
1st quartile (most productive)	2.73	-2.21	1.35
2nd quartile	2.91	-2.35	1.44
3rd quartile	2.91	-2.31	1.44
4th quartile (least productive)	2.86	-2.32	1.42

Notes: We aggregate the sales of all firms that belong to a productivity quartile and calculate their percentage variation with respect to the baseline. CF1 is a case where all the firms belong to the same cultural group. This is, we go from the baseline to $BC\left(z, z^{'}\right) = 1$ for all $z, z^{'}$, which makes the firms to become the closest possible in cultural terms. In this scenario, there are no contracting frictions, as firms know and/or trust each other, and so they pay the minimum trade and matching costs. CF2 is a case where each firm belongs to its own cultural group. Thus, we have a case where $BC\left(z, z^{'}\right) = 0$ for all $z, z^{'}$ and $z \neq z^{'}$, which makes the firms the furthest possible in cultural terms. Under this scenario, firms incur the maximum contracting frictions, for which they pay the maximum trade cost and the maximum matching cost. CF3 is a scenario where trade and matching costs become less sensitive to cultural proximity. In this case parameters β_2 and γ shrink by 50 percent.

2.6.2 Social isolation policies

Since the rise of democracy, efforts have been put in place by the Indian government to end the influence of the caste system in the modern economy (Iyer, Khanna, and Varshney 2013; Munshi 2019). What would have happened if sociopolitical forces perpetuated the social stratification of the caste system? To analyze the maximum impact of social isolation policies we propose case Counterfactual 2 (CF2), where we examine an extreme case in which each firm belongs to its own cultural group. Thus, we have a case where $BC\left(z, z^{'}\right) = 0$ for all $z, z^{'}$ and $z \neq z^{'}$, which makes the firms the furthest possible in cultural terms. Under this scenario, firms incur the maximum contracting frictions, for which they pay the maximum trade cost and the maximum matching cost.

When all firms are the furthest in cultural terms, trade costs and matching costs are the highest. Table 10 presents that in case CF2 total sales fall by 2.23 percent, average intermediate sales go down by 1.20 percent, average intermediate purchases fall by 0.94 percent and prices increase by 1.47 percent. There are also less matches, which is reflected by an average number of buyers that falls by 0.87 percentage points, and an average number of sellers that falls by 0.82 percentage points. As a result, welfare falls by 1.45 percent. Average productivity increases by 0.10 percent, relative to the baseline. Table 11 shows that in case CF2, every firm loses in terms of sales. However, the firms that lose the most are the least productive, which shrinks their weight in the aggregate and, thus, drives average productivity up.

2.6.3 Reducing contracting frictions

Now we turn to study which would be the effect of reducing contracting frictions. For instance, related to our discussion of Section 2.3.1, a policy that improves the quality of courts would reduce the contracting frictions firms face. In terms of our framework, this means that the trade cost and the matching cost become less sensitive to our measure of cultural proximity. Thus, in the Counterfactual 3 (CF3) we analyze a case where parameters β_2 and γ shrink by 50 percent. This captures how reducing contracting frictions affect aggregate outcomes via the channel of trade becoming less reliant on cultural proximity.[25]

Table 10 shows that after reducing contracting frictions in case CF3 the total sales go up by 1.37 percent, average intermediate sales increase by 0.76 percent, average intermediate purchases grow by 0.57 percent and prices fall by 0.87 percent. The number of matches also increases, with the average number of buyers going up by 0.53 percent and the average

[25] Reducing contracting frictions may affect aggregate outcomes through other channels as well, such as more investments in differentiated products, and more trade across longer distances.

number of sellers rising by 0.50 percent. Thus, welfare increases by 0.87 percent. Average productivity goes down by 0.06 percent. In Table 11 we show that in case CF3 all firms gain in terms of sales with respect to the baseline. Nonetheless, it is the lesser productive firms that gain the most, such that their weight in the aggregate increases. This drives the average productivity down.

2.7 Conclusions

We shed light on how cultural proximity shapes the formation of production networks and its implications for welfare. We first provide empirical evidence on the role of cultural proximity for inter-firm trade and the formation of production networks, by leveraging a new dataset of firm-to-firm transactions from a large Indian state, along with data on firm owner names and their cultural proximity derived from India's caste and religious system.

We report three new stylized facts. First, culturally closer firms report higher sales between them. That is, the higher the cultural proximity, the higher the trade in the intensive margin. Second, firms that are culturally closer are more likely to ever trade with each other. This means the higher the cultural proximity, the higher the trade in the extensive margin. Third, firms that are culturally further apart report higher unit prices in their transactions. We show evidence that suggests that the effect we find of cultural proximity on trade is stronger for differentiated goods, which often rely on either formal or informal contract enforcement (Nunn 2007; Rauch 1999). Indeed, we find that differentiated goods are more likely to be produced in and bought by firms that are located in districts with higher contract enforcement, as proxied by court delays. We understand these results as evidence that cultural proximity relates to contract enforcement and trust (Munshi 2014; Munshi 2019).

We build a quantitative general equilibrium model of firm-to-firm trade and cultural proximity. We introduce our measure of cultural proximity as a wedge that affects trade and matching costs, and estimate the key parameters of the model: the semi-elasticity of the trade cost to cultural proximity and the semi-elasticity of matching cost to cultural proximity. We use the model and estimated parameters to quantify the implications for welfare and other model-based statistics of implementing different policies. Welfare increases by 1.76 percent when we evaluate a social inclusion policy, falls by 1.45 percent under social isolation and increases by 0.87 percent when reducing contracting frictions makes firms less reliant on cultural proximity.

In contexts like India, cultural and social networks may be used informally to overcome the lack of formal institutions that uphold contracts. Our paper is among the first to establish the consequences of these cultural ties in the context of trade. We closely study how social relationships influence firm-level decisions and quantify its importance for welfare, both empirical and quantitatively. Our results have strong implications for policy. Promoting social inclusion and mixing via diversity-friendly policies can help facilitate matches and trade, with substantial implications for aggregate output and welfare. Furthermore, investing in reducing contracting frictions will allow firms to not have to rely on cultural ties, and so facilitate matches with more productive and low-cost suppliers, once again improving economic well-being.

Chapter III. Firm-level Elasticities of Substitution and Production Networks

with Devaki Ghose and Gaurav Khanna

3.1 Introduction

The ability of firms to substitute inputs across suppliers is critical for the resilience of supply chains and the transmission of supply shocks. If it is difficult for firms to substitute across suppliers after an adverse supply shock, the shock will amplify by transmitting further downstream through the supply chain. The importance of this mechanism was reflected during the Covid-19 pandemic, where supply chain disruptions drove dramatic reductions in GDP worldwide. For instance, India reported a -7.3% growth rate for the 2020/21 financial year, one of the most significant contractions worldwide and the largest decline in GDP since India's independence.[26] In this paper, we quantify the importance of firm-level elasticities of substitution across suppliers of the same intermediate input to explain large fluctuations in GDP. We provide new estimation strategies and estimates for these elasticities by leveraging regional variation in supply-side shocks induced by the Indian government's massive lockdown policy. We show that this elasticity is key to partly explaining the dramatic decline of the Indian economy during the Covid-19 pandemic. Using new big data computational techniques, we quantify this decline *directly* using information on the economy-wide firm-to-firm network.

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

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

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

We find that inputs within the same HS-4 product category, but across different suppliers are highly complementary. Our estimated elasticity of substitution across suppliers of the same product is 0.55. In various specification tests employing different combinations of fixed effects and different sources of variation, we find that the estimated elasticities lie within a range of 0.49 to 0.65. Our new elasticities show that even within the same HS-4 product category, inputs across firms are highly complementary. The elasticities are similar at the HS-6 product level, and even smaller at the HS-8 product level.[27] As such, even at the very micro level, firm-specific negative shocks contribute to GDP fluctuations. In contrast, Atalay (2017) estimate elasticities at the industry (rather than firm) level. We also estimate the more aggregate firm-level elasticity of substitution across different industries, and find complementarity across industries, in line with Atalay (2017) and Boehm, Flaaen, and Pandalai-Nayar (2019).

As discussed by Taschereau-Dumouchel (2020) and Baqaee and Farhi (2019), the literature so far provides little guidance about estimates of the firm-level elasticity of substitution between suppliers within product categories, even though it is a crucial parameter driving the propagation of shocks. While other work estimates elasticities of substitution across industries (Atalay 2017), across products from different countries (Boehm, Flaaen, and Pandalai-Nayar 2019), or across intermediate goods (Carvalho et al. 2021; Peter and Ruane 2022), such estimates do not yet exist for substitution elasticities across suppliers within the same product category. Estimating elasticities of substitution across different suppliers has been especially challenging for two reasons. First, it is difficult to find detailed information on firm-to-firm transactions with product-specific unit values, reported by each firm. Second, it is challenging to find exogenous sources of variation in firm-level prices

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

(rather than product-level prices) that allows one to credibly estimate these elasticities.

We provide estimates of firm-level elasticities of substitution across suppliers within the same product by leveraging the nationwide, sudden and unprecedented lockdown imposed by the Indian government in March 2020. Importantly, these lockdowns were not homogeneous: districts were categorized into *Green* (mild lockdown), *Orange* (medium lockdown) and *Red* (severe lockdown). Since the lockdowns were sudden and unexpected, they were likely implemented independent of economic fundamentals, and induced strong variation in transactions between firms across India.[28] We use this variation to estimate the firm-level elasticities of substitution across suppliers.

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

In this spirit, we embed our elasticities in a standard network model à la Baqaee and Farhi (2019) augmented with firms and find that the quarterly fall in GDP induced by a negative 25% shock to red-zone firms would be 2.68*pp* less in a model where firms in the same HS-4 product/industry are considered substitutes ($\epsilon = 2$) and 0.99*pp* more when firms in the same HS-4 product are considered almost Leontief ($\epsilon = 0.001$) compared to the baseline case ($\epsilon = 0.55$).[29] In policy counterfactuals, we show that the fall in GDP is much

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

[29] We find that a 25% productivity shock to firms in the red zone reduces GDP by 10.95%. As an empirical

larger if the most connected firms are affected compared to the least connected firms or a random set of firms for a given firm size. The importance of the most connected firms increases non-linearly with the size of the negative productivity shocks and decreases as firms become more and more substitutable. Our experiment suggests that for our baseline value of elasticity of substitution ($\epsilon = 0.55$) and a negative productivity shock of 45% if governments save the better-connected firms, given the same firm sizes, compared to randomly targeting firms, the fall in GDP would be about $0.20pp$ less and $0.31pp$ less compared to targeting the least connected firms. Finally, we quantify how important it is to consider a firm's indirect connectivity in understanding how shocks to the firm can affect aggregate GDP. To be precise, a firm's indirect connections measure not only the number of direct buyers of a supplier but also the buyers' buyers and their buyers, and so on. We find that under our estimated elasticity of ($\epsilon = 0.55$) and a negative productivity shock of 25%, the fall in GDP would be $2.56pp$ less if the government were to bail out firms on the basis of total connectivity as opposed to direct connectivity (counting only the number of direct buyers of a supplier). We see that as the level of the negative productivity shock increases, the difference in aggregate GDP between these two sets of experiments rises, emphasizing the importance of measuring a firm's indirect connections as well.

This paper has three main sections. First, we present reduced-form evidence on the impact of adverse supply shocks on key firm-level variables such as unit values (prices) and the number of transactions (quantities). We leverage the Indian government's sudden lockdown measure that affected firm-to-firm trade across districts, depending on whether firms fall in the *Red* zone (strict lockdown), *Orange* zone (moderate lockdown), or *Green* zone (mostly no lockdown). We find that the prices of intermediate inputs rose during the lockdown, especially if either buyers or sellers were located in *Orange* or *Red* zones. In districts where the seller is in a strict lockdown zone (orange or red), transactions fell dramatically, compared to either the case where the buyer is in a lockdown zone or both are in green zones.

Second, we modify a standard multi-sector firm-level model of input-output linkages by augmenting the production function with substitution across suppliers within the same product category. We derive analytical expressions that relate the relative values of quantities purchased of the same product from different suppliers, to the equilibrium relative prices. That is, within each product category, we quantify how substitutable the different suppliers are. We find that this elasticity of substitution is close to 0.55 for suppliers of an HS-4

benchmark, the state's annual GDP fell by 11.3% in 2020/21.

product, and somewhat smaller for HS-6 and HS-8 products. Thus, following Baqaee and Farhi (2020), after considering second-order effects, adverse firm-level shocks get amplified in the aggregate by propagating through firm-to-firm linkages while positive shocks get dampened. We further explore how these elasticities differ by industry (HS Section), and find that in a handful of industries, suppliers within the same industries are actually substitutes, whereas in others, they are highly complementary. This shows that we should be mindful of heterogeneity across industries in understanding how shocks propagate through supply chains.

Finally, we use the estimated elasticities to analyze how input complementarities at the firm level affect aggregate economic outcomes, and so, how important these complementarities are in explaining GDP declines during the Covid-19 pandemic. We find that a 25% productivity shock relegated only to firms in the red zone reduces overall GDP by 10.96%. This fall would be 2.02*pp* less in a model where firms in the same product category are substitutes ($\epsilon = 1.75$), and 0.75*pp* more when firms in the same HS-4 category are almost Leontief ($\epsilon = 0.001$). Given that the quarterly GDP of this state was close to 32.5 billion USD in 2020, the additional losses due to firm-level complementarities translate into 655 million USD (about 19 USD per capita per quarter), 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 allow large firms (high final sales) or more connected firms (more direct and indirect linkages) to operate. We show that as the level of complementarity and the magnitude of the adverse shock increases, it pays more to save the more connected firms. Much importance, both in policy and academic circles, has been paid to large firms, as Hulten (1978) emphasized the importance of firm sizes in the propagation of shocks through production networks. We show that in the face of large adverse shocks and high levels of complementarity across suppliers, the more connected firms are more important than large firms in shock propagation through the network.

Related work. Our paper connects with two strands of literature. First, we speak to the literature on shock propagation and amplification through supply chains and production networks (Barrot and Sauvagnat 2018; Carvalho et al. 2021; Peter and Ruane 2022; Boehm, Flaaen, and Pandalai-Nayar 2019; Korovkin and Makarin 2020; Ferrari 2022). There are at least three challenges in this literature. First, most firm-to-firm data either do not contain product-level (unit) prices from each supplying firm, or lack the required variation in such prices to estimate firm-level elasticities of substitution across suppliers.[30] Second, and relatedly, limited identifying variation in prices at the buyer-supplier level allows existing work to estimate substitution elasticities across products/industries or across domestic and foreign industries, but not across suppliers within a product category. In contrast, we provide one of the first estimates of the elasticity of substitution across suppliers within a product category: a parameter that is crucial in determining how shocks propagate. Third, the lack of firm-level elasticities across suppliers has so far constrained our assessment of the importance of nodal firms, such as the largest or the most connected firms, in the propagation of shocks through production networks.

We contribute to the literature in each of these dimensions. First, we measure unit prices and quantities at the seller-buyer-product-transaction level. We derive price changes from supply and transportation disruptions in lockdown-affected districts and estimate the firm-level elasticity of substitution between suppliers within a product category. We then quantify this elasticity's importance for amplifying firm-specific supply shocks through a roundabout production network (Baqaee and Farhi 2019). We address previously unanswered questions on the importance of nodal or large firms in shock amplification. We exploit computational innovations in big data to compute the second-order effects of productivity shocks using the entire matrix of production linkages. This innovation helps quantify the non-linear effects of productivity shocks *directly* using the network, without relying on approximations using final sales.[31]

Our paper is also related to research on trade collapses during adverse shocks (Behrens, Corcos, and Mion 2013; Giovanni and Levchenko 2009; Bricongne et al. 2012), and shock transmission through GVCs during Covid, via disruptions to imports/exports or aggregate production (Bonadio et al. 2021; Baqaee and Farhi 2020; Cakmakli et al. 2021; Demir and Javorcik 2020; Gerschel, Martinez, and Mejean 2020; Heise et al. 2020; Lafrogne-Roussier, Martin, and Méjean 2021; Chakrabati, Mahajan, and Tomar 2021). In contrast, we analyze how domestic transactions were affected during Covid lockdowns in a large developing coun-

[30] Carvalho et al. (2021) observe a binary measure of whether firms were connected via buyer-supplier relationships rather than quantities and unit values associated with such transactions. They use a proportionality assumption which precludes estimating the elasticity of substitution across suppliers within a product category, as a buyer sourcing from two suppliers in the same industry will source the same amount given the assumption. Although lacking firm-to-firm price data, Dhyne, Kikkawa, and Magerman (2022) structurally estimate a similar elasticity in the context of imperfect competition models where they restrict the elasticity to be larger than 1 for mark-ups to be relevant.

[31] As firm-to-firm data become common (Panigrahi 2021; Demir et al. 2021; Dhyne, Kikkawa, and Magerman 2022; Alfaro-Urena, Manelici, and Vasquez 2020), our methods can be used to quantify shock propagation through large/complex networks.

try. Our key policy motivation stems from the observation that policymakers worldwide are interested in quantifying the trade-off between strict lockdowns that prevent the spread of the virus but affect GDP through complex buyer-seller networks and more lenient measures that increase production and trade but potentially spread the virus. More importantly, even beyond the immediate Covid crisis, our estimates of how substitutable suppliers are within a product category will help policymakers quantify the economy-wide effects of any disruptive events (e.g., natural disasters or sanctions) on trade and production, that are expected to be reasonable short lived.[32]

3.2 Data and Context

Firm-to-firm trade. Our primary data source is daily establishment-level transactions with distinct information on establishment locations.[33] This data is provided by the tax authority of a large Indian state with a diversified production structure, roughly 50% urbanization rates, and high levels of population density. To compare its size in terms of standard firm-to-firm transaction datasets, the population of this Indian state is roughly three times the population of Belgium, seven times the population of Costa Rica, and two times the population of Chile.

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

Each transaction reports a unique tax code identifier for both the selling and buying establishments, all the items contained within the transaction, the value of the whole trans-

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

[33]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.

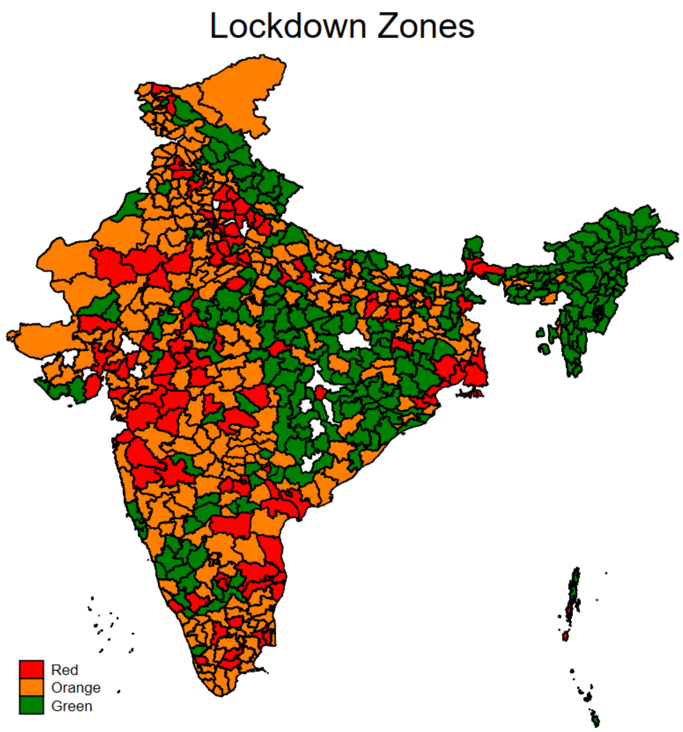
action, the value of the items being traded up to 8-digit HS codes,[34] quantity of each item, units, and mode of transportation. Each transaction also reports the ZIP code of the selling and buying firms, which we use to merge with other district-level data.

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

Lockdowns. On March 25th 2020, India unexpectedly imposed strict lockdown policies nationwide. The designated severity of the lockdown varied by districts, and was implemented nationwide at the district level, where each district was classified between *Red*, *Orange*, and *Green* zones according to the severity of Covid cases in each district. Yet, at that time, there were barely any Covid cases in India, as the entire country averaged about 50 cases a day (as opposed to about 400,000 cases a day the following year).

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

Figure 12: India's lockdown zones in March, 2020



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

In Figure 12, we map the distribution of lockdowns across India. Districts in the red zone saw the strictest lockdowns, with rickshaws, taxis and cabs, public transport, and barbers/spas/salons remaining shut. E-commerce was allowed for essential services. Orange and green zone districts saw fewer restrictions. Orange zones allowed the operation of taxis and cab aggregators, as well as the inter-district movement of individuals and vehicles for permitted activities. In addition to the activities allowed in orange zones, buses were allowed to operate with up to 50% seating capacity and bus depots with 50% capacity in green zones.[35]

Throughout the paper, we use this color scheme as the treatment across Indian districts. In particular, each firm is located within a district, so treated firms are located within a *Red*, *Orange*, or *Green* district between March and May 2020.

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

Physical and cultural distance. We use different measures of *distance* which we include as controls in our empirical results. The measures of geographic distance between districts calculate the length of the shortest distance between district centers. The measure of linguistic distance between Indian districts is from Kone et al. (2018) who using 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.

Other 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`. For each firm, we construct two variables that measure the firm's exposure to global demand and supply shocks that vary at the product and country level, following Hummels et al. (2014). The construction of these exposure variables are described in detail in online data Appendix C.3.

Summary statistics. We present some key summary statistics from the administrative trade data in Table C1. 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.

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.

To further understand the composition of economic activity of the Indian state of our analysis, in Table C2 we show what types of goods firms within the state sell and buy, and to which destinations. In out state, firms are mostly in the business of selling vegetables, plastics, and minerals; and of buying machinery, metals, and vegetables. In terms of the type of trade, firms in our state mostly sell to firms in other Indian states. This contrasts with how firms in our state buy intermediates, where the share of purchases that come from within the state is almost the same as from other Indian states. Finally, international exports and imports represent a non-negligible but rather small share of both sells and purchases.

Before using the lockdown variation to understand how firm to firm transactions are affected, we verify the stringency of these lockdowns in Figure C3 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 3 – Feb 6, 2020.[36] 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 (time spent in retail and recreation, grocery and pharmacy, parks, commuting, and workplaces) in red zones compared to green zones; with orange zones in between. People in red zones also 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 small.

3.3 Reduced-Form Evidence

In this section we outline a simple empirical specification to provide evidence showing the role of lockdown policies on key outcome variables for firm-to-firm trade. We show that the sudden Covid-19 lockdown policies between March and May 2020 led to a rise in unit values, and a fall in the monthly number of transactions between firms.[37] In subsequent sections, we exploit this variation to estimate firm-level elasticities of substitution across intermediate inputs.

3.3.1 Empirical specifications

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

Seller-level regressions. We estimate the following specification:

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

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

$$Y_{si,t} = \iota_{i,o_{(s)}} + \iota_{i,t} + \sum_{t\neq -1} \beta_t Red_{o_{(s)}} + \sum_{t\neq -1} \gamma_t Orange_{o_{(s)}} + X\delta + \epsilon_{si,t}, \tag{3.1}$$

where $Y_{si,t}$ are either unit values or the log number of transactions for seller s in HS-4 product i in month t, $\iota_{i,t}$ are 4-digit HS-by-month fixed effects, $\iota_{i,o_{(s)}}$ are product-by-district fixed effects (i.e. fixed effects based on the district o where seller s resides). We extend the analysis to the HS-6 and HS-8 level in the appendix. X are controls that include number of Covid cases, deaths, and recoveries, and exposure to international demand and supply shocks as discussed in Appendix C.3. We control for the Covid cases and deaths since these are the variables on which the government based its lockdown decisions (Ravindran and Shah 2020). The covariates of interest are $Red_{o_{(s)}}$ and $Orange_{o_{(s)}}$. The first one is an indicator 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 estimates of interest are β_t and γ_t. Our base time category is February 2020 which is just before lockdowns began. Standard errors are clustered at the seller's origin district level.

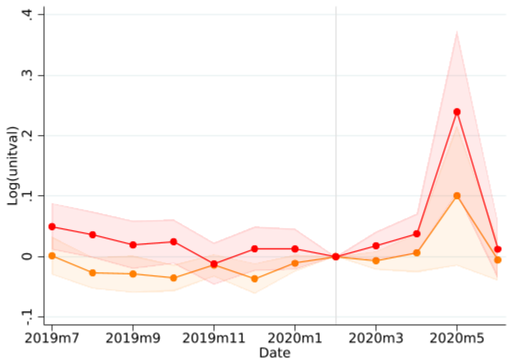
Seller/buyer-level regressions. At the seller-buyer level we estimate the specification:

$$Y_{si,b,t} = \sum_{(x,z)\in\Omega} \sum_{t\neq -1} \beta_t^{xz} \left(\gamma_{o_{(s)}}^{x} \times \gamma_{d_{(b)}}^{z} \right) + \delta_{o_{(s)}} + \delta_{d_{(b)}} + \delta_{i,t} + \beta_1 \log \ dist_{od} + X\delta + \epsilon_{si,b,t} \ , \tag{3.2}$$

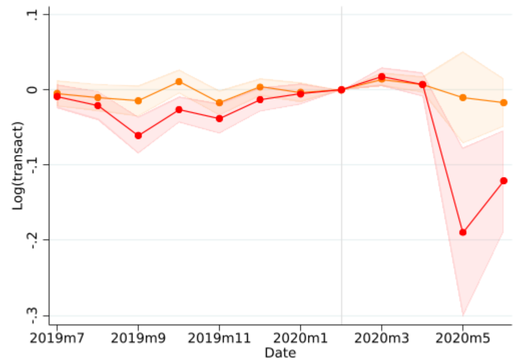
where $Y_{si,b,t}$ are unit values or number of transactions in logs between seller s in HS-4 product i and a buyer b in month t. $\delta_{o_{(s)}}$, $\delta_{d_{(b)}}$, and $\delta_{i,t}$ are origin, destination, product-by-month fixed effects. $dist_{od}$ is a vector of cultural and geographic distance variables, and 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 contains our estimates 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 (3.2). $\gamma_{o_{(s)}}^{x}$ 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 estimates of interest are β_t^{xz}. Our base time category is February 2020 which is just before lockdowns began. Standard errors are two-way clustered at the origin and destination district level.

Figure 13: Seller-level reduced-form event studies

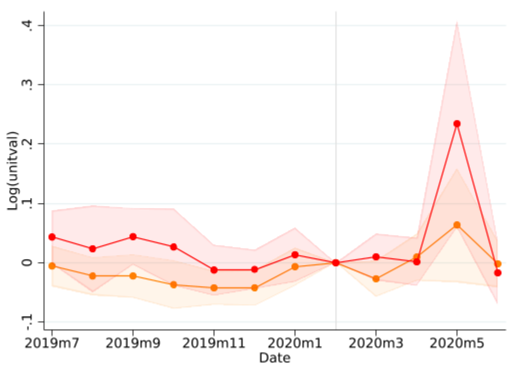
(a) Unit value, 4-digit HS



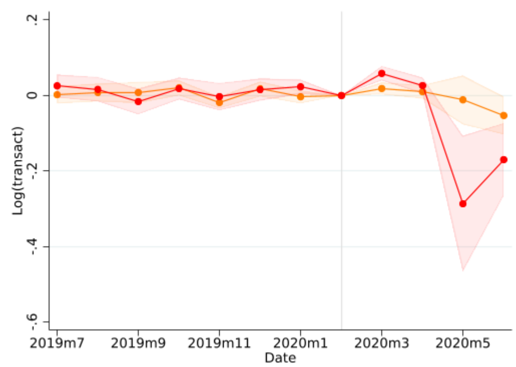
(b) # Transactions, 4-digit HS



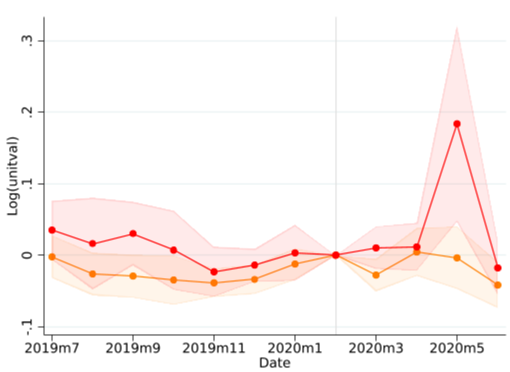
(c) Unit value, 8-digit HS



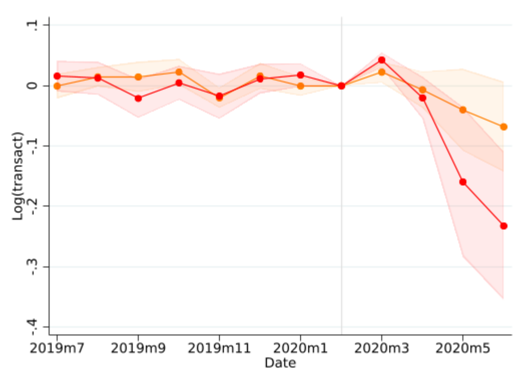
(d) # Transactions, 8-digit HS



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



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



Notes: This figure is comprised of 6 plots. Each plot shows estimates for β_t and γ_t from Equation (3.1). The values of the estimates are all in comparison to sellers in *Green* districts in February 2020. The dependent variable on the left side is in log unit values; on the right side, in log number of transactions. Each row varies by the definition of a product-group, and the fixed effects included in the regression. In the first row, a product is 4-digit HS codes and fixed effects HS/month and district. In the second row, a product is an 8-digit HS code and fixed effects HS/month and district. In the third row, a product is an 8-digit HS code and fixed effects HS/month and district/HS. Standard errors are clustered at the district level. All controls mentioned in the paper are included. The shaded area are confidence intervals.

3.3.2 Reduced-Form Facts

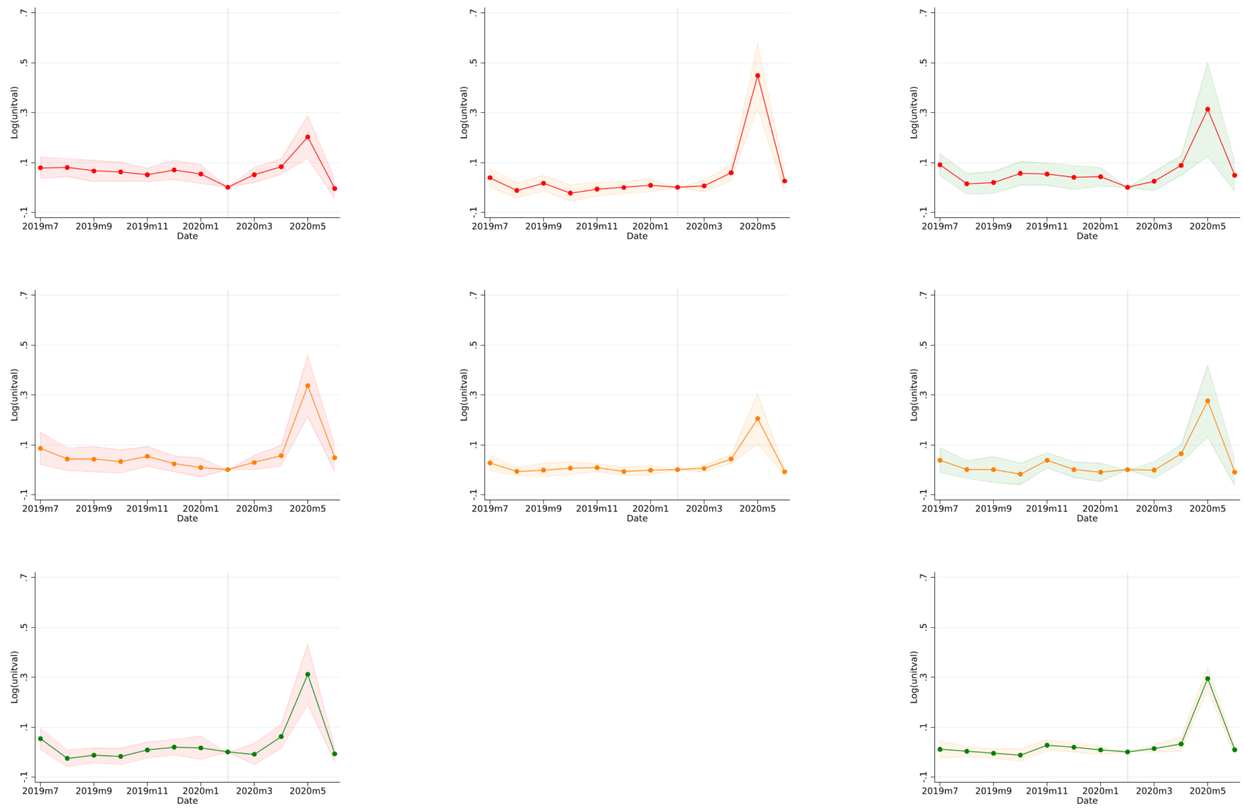
In this section we present two facts from the specifications we laid out in the previous section.

Fact 1: Sellers' unit values disproportionately rose and trade fell in more severe lockdown zones. The first two panels of Figure 13 plot the coefficients β_t and γ_t from Equation (3.1), representing changes in log unit values and log number of transactions with respect to *Green* districts in February 2020 (the base category). In May 2020, sellers' unit values in *Red* districts rose by 25pp, and in *Orange* districts rose by around 10pp with respect to the base category.

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

The middle two panels of Figure 13 repeats the same exercise with a finer product definition, using 8-digit HS codes. Results remain virtually the same. In the last row of Figure 13 we include a stronger set of fixed effects (e.g., district-by-product), and results remain the same.

Figure 14: Unit Value, Seller-Buyer Level Regressions



Notes: In each plot, the horizontal axis is the month, and the vertical one is the estimate of interest associated with log unit values as in Equation (3.2) for each month. Regressions include product-by-month, origin district, and destination district fixed effects. Standard errors are two-way clustered at the origin and destination state level. A product is a 4-digit HS code. All controls mentioned in the paper are included. The vertical line in January 2020 splits pre and post-lockdown 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.

Fact 2: Equilibrium unit values rose and number of transactions fell in more severe lockdown zones. We now report the results from our seller/buyer-level specification. In Figures 14 and 15 we report the estimates for β_t^{xz} in Equation (3.2), where the estimates are in comparison to cases when both sellers and buyers were located in Green districts in February 2020.

In the first row of Figure 14 we plot the coefficients from regression (3.2) where the seller is in the red zone, and the buyer is in red, orange, and green zones respectively. Similarly, in the second row of Figure 14, we plot the coefficients from regression (3.2) where the seller is in the orange zone, and in the third row, we plot the coefficients from regression (3.2) where the seller is in the green zone (and the buyer is in red and orange zones respectively).

Figure 15: Number of Transactions, Seller-Buyer Level Regressions

Notes: In each plot, the horizontal axis is the month, and the vertical one is the estimate of interest associated to log number of transactions as in Equation (3.2) for each month. Regressions include product-by-month, origin district, and destination district fixed effects. Standard errors are two-way clustered at the origin and destination state level. A product is a 4-digit HS code. All controls mentioned in the paper are included. The vertical line in January 2020 splits pre and post-lockdown 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.

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

Our two facts jointly imply that prices where either seller or buyers were located in red districts were higher during the lockdown in comparison to districts where the lockdowns were mild (green zones). This suggests that the lockdown indeed induced variation in prices that we will later leverage to estimate elasticities of substitution across intermediates.

3.4 Model

We build a quantitative general equilibrium model of firm-to-firm trade based on Baqaee and Farhi (2019), where the productive sector is perfectly competitive.[38] We adapt the general nested CES structure to reflect the possibility that suppliers within the same product category could be substitutes or complements, derive estimating equations, and use the model to simulate the effects of negative productivity shocks on GDP. Firms combine inputs in a CES fashion under three tiers. In the first tier, firms combine labor and aggregated intermediates. In the second tier, aggregated intermediates are a combination of intermediates by product composites. In the third tier, product composites are constructed by suppliers of intermediates.

There are N firms producing N goods using the production function

$$y_{nj} = A_n \left(w_{nl} \left(l_n \right)^{\frac{\alpha-1}{\alpha}} + \left(1 - w_{nl}\right) \left(x_{nj} \right)^{\frac{\alpha-1}{\alpha}} \right)^{\frac{\alpha}{\alpha-1}} , \tag{3.3}$$

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

$$x_{nj} = \left(\sum_{i=1}^{I} w_{i,nj}^{\frac{1}{\zeta}} \left(x_{i,nj} \right)^{\frac{\zeta-1}{\zeta}} \right)^{\frac{\zeta}{\zeta-1}} , \tag{3.4}$$

where ζ is the elasticity of substitution between inputs from different product categories, and $w_{i,nj}$ is the importance of inputs of product category i for buyer b of product j. $x_{i,nj}$ are intermediate inputs from product i going to firm n producing product j,[39] which are in turns constructed as:

$$x_{i,nj} = \left(\sum_{m=1}^{N_i} \mu_{mi,nj}^{\frac{1}{\epsilon}} x_{mi,nj}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}} , \tag{3.5}$$

where $x_{mi,nj}$ are intermediate inputs from firm m of product i sold to firm n producing prod-

[38] We do not rely on models featuring market power (Edmond, Midrigan, and Xu 2018; Alviarez et al. 2021) since the evidence from the data suggests that the market structure in this Indian state is closer to perfect competition. The median HHI across 4-digit HS product categories is 0.1041, which implies a low level of market concentration within a product category.

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

uct j, and $\mu_{mi,nj}$ is the importance of input from supplier m of product i in the production of buyer n of product j. We consider a fixed set of firms F and product categories I, where $N = |F|$ is the total number of firms in the economy, and N_i is the number of firms producing product i. ϵ is the elasticity of substitution across suppliers within the same product category. The above production functions work for reproducible factors. For non-reproducible factors, in our case labor, the production function is an endowment: $Y_f = 1$.

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

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

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

Model in *standard-form.* To write the economy in standard form 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 corresponds 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 $\widehat{\theta} = (\sigma, \alpha, \zeta, \epsilon)$.

Formally, a nested-CES economy in standard form is defined by $\left(\widehat{\Omega}, \widehat{\theta}\right)$. 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 (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 is given by

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

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:

$$\frac{d^2logY}{dlogA_j dlogA_i} = \frac{d\lambda_i}{dlogA_j} = \sum_k (\theta_k - 1)\lambda_k Cov_{\Omega_{(k)}}(\Psi_{(i)}, \Psi_{(j)}). \tag{3.6}$$

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 product category j, all products 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 product i depends on the size of product k as captured by its *Domar* weight λ_k, the elasticity of substitution θ_k in k's production function, and the extent to which the supply chains that connect i and j to k coincide with one another, as given by the covariance term.

3.4.1 Equations to estimate firm-level elasticity of substitution across suppliers

Using the model outlined above, in this section we derive the firm-level elasticity of substitution across suppliers within a product. We introduce a notation change to facilitate the exposition: a firm n can be either a buyer $b \in F$ or a seller $s \in F$. A firm b with product $j \in I$ maximizes profits subject to its technology and to a CES bundle of intermediate inputs:

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

subject to (3.3), (3.4), and (3.5). ϵ from Equation (3.5) is the elasticity of substitution across different suppliers within the same product category. 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 product. Details about the optimization problem are in Appendix C.4.1. The maximization problem yields the following expression:

$$\log\left(\frac{PM_{si,bj}}{PM_{i,bj}}\right) = (1-\epsilon)\log\left(\frac{p_{si,bj}}{p_{i,bj}}\right) + \log\left(\mu_{si,bj}\right), \tag{3.7}$$

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 3.5.

3.4.2 Equations to estimate firm-level elasticity of substitution across products

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

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

subject to (3.3) , (3.4), and $p_{i,bj} = \left(\sum_s \mu_{si,bj} p_{si,bj}^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}$. ζ from Equation (3.4) is the firm-level elasticity of substitution across products i we estimate. Notice that in this case, we need values for ϵ and $\mu_{si,bj}$ to calculate prices. We consider $\epsilon = \hat{\epsilon}$, where $\hat{\epsilon}$ is our estimate, and we recover $\mu_{si,bj}$. Details on the optimization problem are in Appendix C.4.2. The maximization problem yields the following expression:

$$\log\left(\frac{PM_{i,bj}}{PM_{bj}}\right) = (1-\zeta)\log\left(\frac{p_{i,bj}}{p_{bj}}\right) + \log\left(w_{i,bj}\right), \tag{3.8}$$

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 estimating equation for the firm-level elasticity of substitution ζ which we take to the data, as described in Section 3.5.

3.5 Estimation

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

3.5.1 Estimating equations for ϵ and ζ

In order to estimate ϵ from Equation (3.7), the first major challenge we face is that the price index $p_{i,bj}$ includes the unobserved quantity $\mu_{si,bj}$ which denotes the importance of input from supplier s of product i in the production of buyer b producing product j. This unobserved quantity could depend on a number of factors such as unobserved input demand shocks or the buyer's preference for certain inputs. In order to construct changes in price indices that are observable, we follow Redding and Weinstein (2020) in assuming that the overall importance of a product in a buyer's input use does not change between two consecutive months, even though the importance of inputs from suppliers within a product category can change.[40] We arrive at the equation below that links the overall expenditure share on a certain supplier s (as a share of total expenditure on product i) to the corresponding relative prices:

$$\log\left(\frac{\widehat{PM}_{si,bj,t}}{\widehat{PM}_{i,bj,t}}\right) = \omega_{b,t}+\omega_{i,t}+\omega_{b,i}+\omega_{s,i}+(1-\epsilon)\log\left(\frac{\widehat{p}_{si,bj,t}}{\widehat{\widetilde{p}}_{i,bj,t}}\right)+\log\left(\widehat{\lambda}_{i,bj,t}\widehat{\widetilde{s}}^{*}_{i,bj,t}\right)+X\beta+\xi_{si,bj,t}\ , \tag{3.9}$$

where $\widehat{x}_t = \frac{x_t}{x_{t-1}}$ are variables in changes with respect to the previous month. $[\omega_{b,t}, \omega_{i,t}, \omega_{b,i}, \omega_{s,i}]$ is a set of fixed effects, including buyer-by-month, product-by-month, buyer-by-product, and seller-by-product fixed effects. $\widetilde{p}_{i,bj,t} = \prod_{s\in\Omega^*_{i,bj,t}} p_{si,bj,t}^{\overline{\frac{1}{N^*_{i,bj,t}}}}$ is a geometric mean of unit values across common suppliers, where $\Omega^*_{i,bj,t} \equiv \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$ is the set of common suppliers for buyer b that appear in both the current and previous month, and $N^*_{i,bj,t} \equiv \Omega^*_{i,bj,t}$ is the number of common suppliers for buyer b in month t. X are controls, including exposure to foreign demand and supply shocks, the number and severity of Covid cases, and geographic and cultural distance.

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

[40]This assumption simply requires that, for instance, a shoemaker's overall preference for leather in shoe-manufacturing does not change, although its preference for leather from certain suppliers can change. That is, demand-shocks may change $\mu_{si,bj,t}$ (e.g., the demand for leather from certain suppliers), but the geometric mean of $\mu_{si,bj,t}$ across suppliers within a product is stable between t and $t-1$. This enables us to construct changes in price indices that are not dependent on $\mu_{si,bj,t}$, but are directly observed in the data (details in Appendix C.4.1).

induced variation in relative prices.

Our setup has the advantage that we can decompose the change in price buyer b pays for inputs from seller s between $\widetilde{p}_{i,bj,t}$, the change in expenditure share $\widehat{\widetilde{s}}^{*}_{i,bj,t}$ and a Feenstra (1994) correction term $\widehat{\lambda}_{i,bj,t}$ that takes into account the fact that sellers enter and exit in the data. More details are in Appendix C.4.1. Standard errors are two-way clustered at the origin and destination state level.

Now, to estimate ζ from Equation (3.8), there are two issues to address. First, notice that the price index $p_{i,bj}$ is a function of (unobservable) demand shocks $\mu_{si,bj,t}$, and ϵ. Second, the price index $p_{bj,t}$ is also a function of unobservable product-level demand shocks $w_{i,bj,t}$, which makes their computation challenging.

First, we construct price indices as $p_{i,bj,t} \equiv \left(\sum_s \mu_{si,bj,t} p_{si,bj,t}^{1-\widehat{\epsilon}}\right)^{\frac{1}{1-\widehat{\epsilon}}}$, where $\widehat{\epsilon}$ are estimated previously, $p_{si,bj,t}$ come directly from the data, and demand shocks $\mu_{si,bj,t}$ are constructed recursively. This recursive construction of demand shocks come from predicting residuals from Equation (3.9) and setting an initial value for shocks $\mu_{si,bj,0}$ (Appendix C.4.2).

Second, we construct buyer-level price indices $p_{bj,t}$ following Redding and Weinstein (2020). We assume that the overall importance of the composite intermediates at HS-4 level in the production function does not change between consecutive months. As such, we can construct this price independent of product-level demand shocks $w_{i,bj,t}$ after controlling for buyers' expenditure shares by product. More details about this are in Appendix C.4.2. We then derive the following expression we take directly to the data:

$$\log\left(\frac{\widehat{PM}_{i,bj,t}}{\widehat{PM}_{bj,t}}\right) = \omega_{b,t} + \omega_{i,t} + \omega_{b,i} + (1-\zeta)\log\left(\frac{\widehat{p}_{i,bj,t}}{\widehat{\widetilde{p}}_{bj,t}}\right) + \log\left(\widetilde{s}_{bj,t}\right) + \xi_{i,bj,t} \ , \tag{3.10}$$

where $[\omega_{b,t}, \omega_{i,t}, \omega_{b,i}]$ are a set of buyer-by-month, product-by-month, and buyer-by-product fixed effects, which again account for a wide array of demand, product shocks, and buyer-product specific charachteristics. $\widetilde{p}_{bj,t} \equiv \prod_{i=1}^{N_{bj,t}} \widetilde{p}_{i,bj,t}^{\frac{1}{N_{bj,t}}}$ is the geometric mean of unit values across products that buyer b purchases, and $\widetilde{s}_{bj,t} \equiv \prod_{i=1}^{N_{bj,t}} \widetilde{s}_{i,bj,t}^{\frac{1}{N_{bj,t}}}$ is the geometric mean of expenditure shares across products. Detailed derivations are in Appendix C.4.2.

3.5.2 Addressing endogeneity concerns

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

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

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

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

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

Seller-level instruments. We need supply-side shifters to obtain unbiased elasticities of substitution. In that sense, shocks induced by the Covid-19 lockdown policies that only impact sellers would provide that variation. In Equation (3.11) below we formalize this intuition, so

$$log(\widehat{p}_{si,bj,t}) = \beta^R Red_{o_{(s)}} Lock_t + \beta^O Orange_{o_{(s)}} Lock_t + \epsilon^{\nu}_{si,bj,t} \ , \tag{3.11}$$

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

Seller/buyer-level instruments. The transportation of supplies from the location of the supplier to the buyer 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. We allow trade cost to change over time such that we can leverage the Covid-19 lockdown policy. In particular, we assume

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

After considering this functional form for trade costs into the expression of prices and log-differencing, we obtain

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

We leverage the Covid-19 lockdown as an exogenous shifter that only influences travel time between locations of seller s and buyer b, as reflected in Equation (3.12) below.

$$log(\widehat{p}_{si,bj,t}) = \beta^R Red_{o_{(s)}d_{(b)}} Lock_t + \beta^O Orange_{o_{(s)}d_{(b)}} Lock_t + \epsilon^{\nu}_{si,bj,t}. \tag{3.12}$$

Detailed derivations are in Appendix C.4.1. $Red_{o_{(s)}d_{(b)}}$ and $Orange_{o_{(s)}d_{(b)}}$ are the share of districts designated as *Red* and *Orange*, respectively, along the route between seller s and buyer b. We constructed these variables using Dijkstra algorithm for least-cost routes. Details about the implementation of this algorithm are in Appendix C.3.

Finally, we also instrument the changes in relative prices in Equation (3.10) to estimate ζ. We do this because of potential unobservable product-level demand shocks that also induce an upward bias to estimates of ζ. To construct our instruments, we leverage the seller-level and seller/buyer-level instruments we used to estimate ϵ and calculate weighted averages across suppliers to instrument on the change of relative prices for buyers. The intuition is that buyers that purchased inputs either from a larger share of sellers in *Red* zones, or from sellers located in districts where the route is comprised or a larger share of *Red* zones were more exposed to Covid-19 lockdowns. More details are in Appendix C.4.2.

Discussion of instruments. The instruments induce buyers of certain types to be more affected than others based on their production networks. The Local Average Treatment Effect (LATE) may not represent the Average Treatment Effect (ATE) if buyers in *Red*, *Orange*, and *Green* zones already traded intensively with sellers in certain lockdown zones, and there is heterogeneity in responses. For instance, if buyers in *Red* traded mostly with sellers in *Red*, then our instrument may estimate effects on firms induced by having more *Red* sellers, and so it would upweight effects on buyers in *Red*. In Figure C1 we run two sets of balance check to investigate these patterns. These checks show that, in general, sellers from *Red*, *Orange*, and *Green* zones had similar interactions with buyers from *Red*, *Orange*, and *Green* zones.

We also consider whether certain products are sourced intensively from firms located in certain zones. For instance, if all the rubber supply of firms in this production network comes from suppliers in *Red* zones, then buyers of rubber would find it increasingly difficult to find suppliers. Once again, if there is heterogeneity in responses by product category, our estimated LATE elasticity would weigh rubber products higher than non-rubber products. While not a source of bias, it does affect the interpretation of the estimated parameter. In Figures C3a and C3b, we plot the shares of total purchases of each industry (HS Section) that are sourced from firms in *Red*, *Orange*, and *Green* zones. With the exception of the small HS industry 19 (arms and ammunitions), there is no noticeable degree of concentration of suppliers from any particular zone.

3.5.3 Estimation results: Firm-level elasticities of substitution across suppliers

First, we report OLS estimates in Table 12. The implied elasticities exhibit a robust value of 0.78 across all the different specifications. In column (1), we include both buyer/month and HS/month fixed effects. In column (2) we also include buyer/HS and seller/HS fixed effects. We obtain a similar elasticity of 0.77. To test whether our estimates vary by product aggregation, in columns (3) and (4) the estimations are based on 6-digit and 8-digit HS codes. The elasticities are around 0.75, so the estimates do not significantly change. Since these elasticities are below 1, these estimates suggest that, at the firm level, suppliers act as complements rather than substitutes for buyers. This is important for aggregate incomes since, from Equation (3.6) we can see that, once we take into account second order effects, an elasticity of substitution less than 1 implies that the aggregate impacts of negative shocks are amplified.

Table 12: OLS, firm-level elasticity of substitution across suppliers

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

Notes: OLS estimates from Equation (3.9). The first row reports the estimates associated with changes in relative unit values in logs. Standard errors are two-way clustered at the origin and destination state level, and are reported in parentheses below each estimate. The fifth row reports the implied value for ϵ, which is 1 minus the estimate on the first row. The table contains four columns. Each column correspond to different specifications on how we define a product (4-digit, 6-digit, or 8-digit HS codes) and of fixed effects, as pointed out by the last five rows of the table. All specifications include the controls mentioned in the paper.

Nevertheless, as we describe in the previous section, it is possible that OLS estimates are contaminated by simultaneous demand shocks that happened during Covid-19. In Table 13 we report 2SLS estimates based on our proposed instruments. We find evidence that inputs across different suppliers of a firm within the same 4-digit HS product category are highly complementary, ranging from $0.49 - 0.65$, depending on the set of fixed effects and instruments we use. Later we show similar patterns for HS-6 and HS-8 categories. Our preferred specification is column (3) with an elasticity of 0.55, where we use both the seller and the seller-buyer level instrument, essentially deriving variation from both sellers' production costs and transportation costs. We include buyer/month and HS/month fixed effects that account for time-varying demand shocks, and also account for entry/exit with the Feenstra (1994) term. Each specification reports a high Kleibergen-Paap F-statistic, indicating that our instruments are statistically relevant. In columns (1) and (2) we use the seller-level and seller/buyer-level instruments separately. The elasticities are 0.49 and 0.6 respectively, which also reflect complementarity. Finally, in column (4) we also include buyer/HS and seller/HS fixed effects, and the elasticity rises to 0.66.

Table 13: 2SLS, firm-level elasticity of substitution across suppliers

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

Notes: IV-2SLS estimates from Equation (3.9). The set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$. That is, a supplier s of buyer b is considered *common* if they also traded during the previous month. The first stage uses either bilateral or seller-level instruments, as pointed out by rows six and seven. Bilateral instruments correspond to Equation (3.12), while seller-level instruments correspond to Equation (3.11). The first row reports estimates associated with changes in relative unit values in logs. Standard errors are two-way clustered at the origin and destination state level, and are reported in parentheses below each estimate. The fourth row reports the Kleibergen-Paap F statistic from the first stage. The fifth row reports the implied value for ϵ, which is 1 minus the estimate on the first row. A product is a 4-digit HS code and the treatment period is March-May 2020. The table contains four columns. Each column corresponds to different combinations of instruments and of fixed effects, as pointed out by the last six rows. All specifications include the controls mentioned in the paper.

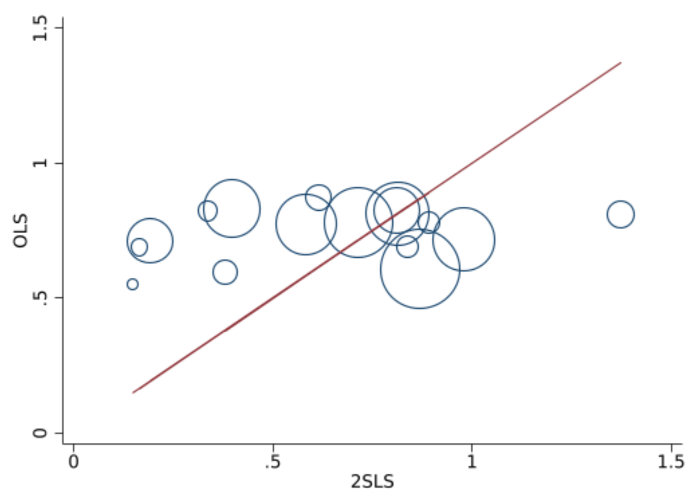
The IV estimates for ϵ are smaller than the OLS estimates. As discussed in Section 3.5.2, the bias is in the expected direction if we expect the Covid-19 shock to also induce negative demand shocks, thereby biasing up OLS estimates of ϵ. We may expect that our estimated elasticity be lower for the sub-sample of buyers who did not have more than one supplier to source inputs from. In Table C3, we restrict our sample to cases when a buyer traded with at-least two sellers in two consecutive periods. Column (3), our preferred specification, yields an elasticity of substitution of 0.58, very close to the estimate from our main specification.

To examine differences by the level of aggregation of the product, we rerun our main specification in Table C4 using HS-6 and HS-8 as product definitions. Finer product classifications (e.g., HS-8) may imply that there are fewer suppliers one may be able to source from, and so we may expect a lower elasticity of substitution between suppliers. In columns (1) and (3) we replicate our main specifications, with elasticities of 0.43 (for HS-6) and 0.06

(for HS-8) respectively. These numbers reflect even higher degrees of complementarity when we consider a more granular notion of product. Overall, these patterns suggest that inputs are highly specific for buying firms.

Elasticity heterogeneity by Industry. We now analyze whether the degree of substitution across suppliers varies by industry (HS section). The idea is that firms that source from highly specific intermediate inputs (i.e. processed foods) should report a lower elasticity of substitution across suppliers than firms that source from more general inputs (e.g. textiles). In Table C5 and Figure 16 we show the estimates of this elasticity of substitution across suppliers by twenty one broad industries. We find that the OLS elasticity of substitution across suppliers by industry lies in the range of 0.7-0.9. Once we instrument for the unit values with the Covid-19-induced lockdown variation, we find that there is wider heterogeneity across industries. Indeed, we find that that *Processed foods* yield an elasticity of 0.19, while *Textiles* yield an elasticity of 0.81. Also, while for the majority of the industries we find evidence for complementarity, there are some industries such as Plastics, Vegetables, and Handicrafts where suppliers within a HS-4 product are likely substitutes.

Figure 16: Elasticities ϵ by seller's 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 IV-2SLS. An industry is an HS section. The size of each bubble is determined by total sales in the corresponding industry. See Table C5 for industry-specific numbers.

3.5.4 Estimation results: Firm-level elasticities of substitution across products

In Table 14, we report our estimates for the firm-level elasticity of substitution across products. In column (1) we show the OLS estimate of $\zeta = 0.91$, which reflects complementarity between product categories. Columns (2) and (3) show cases when we define products more granularly. In this case, the elasticities are around 0.8, which also reflects complementarity between products.

Table 14: Firm-level elasticity of substitution across products

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

Notes: IV-2SLS estimates from Equation (3.8). Price indices are constructed by recovering the residuals used in the corresponding specification when estimating ϵ and the corresponding estimate of ϵ. The first three columns are OLS estimates of ζ; the last three, 2SLS of estimates of ζ using both weighted averages of both bilateral or seller-level instruments across sellers. Bilateral instruments correspond to Equation (3.12), while seller-level instruments correspond to Equation (3.11). Each column corresponds to a different combination of fixed effects and definition of product. Columns (1)-(2) and (4)-(5) correspond to our preferred specification when estimating ϵ and 4-digit and 8-digit HS codes. In columns (3) and (6) we also include buyer/HS fixed effects. The first row reports the estimates associated with changes in relative unit values in logs. Standard errors are clustered at the buyer's district level, and are reported in parentheses below each estimate. The fourth row reports the Kleibergen-Paap F statistic from the first stage. The fifth row reports the implied value for ϵ, which is 1 minus the estimate on the first row. The sixth row denotes whether estimators are OLS or 2SLS. The sixth row mentions the definition of product. The last three rows indicate the combination of fixed effects.

In columns (4)-(6) we report our estimates of ζ under 2SLS estimation after using a weighted average of instruments across buyers' sellers as discussed in Section 3.5.2. Our specification in column (4) reports a value of 0.68, reflecting that simultaneous negative demand and supply shocks during Covid-19 led to an underestimation of ζ under OLS. This elasticity is higher than the 2SLS elasticity of substitution across suppliers for the same product ($\epsilon = 0.55$), reflecting a lower degree of complementarity across products compared to suppliers.[41] In columns (5) and (6), similar values for this elasticity hold when we define a product as 8-digit HS codes, and after the inclusion of buyer/HS fixed effects. Finally,

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

first stage F-stats are high, which reflects the statistical relevance of our weighted averaged instruments.

Unlike the elasticity of substitution across suppliers within a product category, there have been previous attempts in the literature to estimate the elasticity of substitution across products or industries. In particular, other work has estimated a wide range of values for parameters akin to ζ depending on the aggregation of the industry and on the research question. Our elasticity is close to Boehm, Flaaen, and Pandalai-Nayar (2019) who estimate an elasticity across HS-10 products that lies between $0.42 - 0.62$ for non-Japanese affiliates and 0.2 for Japanese affiliates. Atalay (2017) finds an estimate of around 0.1 for 30 aggregated industries using US data.

3.6 Quantification and Counterfactuals

In this section, we use both data from our production network and our newly estimated elasticities to quantify the role of these elasticities in the propagation of shocks. To do this, we need to write down the Leontief matrix in standard form. Given the production structure of our economy, we need four submatrices: (i) firm purchases of 4-digit HS products, (ii) firm sales of 4-digit HS products, (iii) labor employed by each firm, and (iv) final sales by each firm. The first two submatrices are directly constructed from the firm-to-firm trade data from the pre-Covid period of March 2019 to February 2020. Labor employed and final sales by firms are obtained by merging in firm-level data from Indiamart, which contains information on firm-level employment and final sales.[42] For more details for this, see Appendix C.3.

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

For more details on the derivation of the shock propagation equation and its numerical implementation, see details in Appendix C.5. While previous work also quantify the effect of firm-level shocks on aggregate GDP, they mostly rely on changes in firm-level final sales rather than the direct production network. Using the production network directly, exponentially increases computational complexity from the order of N to $(N+I)\times(N+I)$, where N is the

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

no of firms, and I is the no of distinct products. As such, we use computational innovations in big data to implement this procedure.

Note that our quantification exercises in this section are conditional on the products that firms buys/sells being given at the extensive margin, even though a firm can change its set of buyers/suppliers, as documented by Khanna, Morales, and Pandalai-Nayar (2022). We therefore need to empirically assess whether the set of HS-4 products a buyer buys and the set of HS-4 products that a seller sells, changes between the pre and the post Covid period. We do this by inspecting whether both sellers and buyers of each product continued to trade in their corresponding product categories after Covid-19 lockdowns. In Figure C4 we show the product-level distribution of share of sellers that sold and buyers that purchased goods of that product during both time periods t and $t-1$, where t is a 6-month window before and after the lockdowns.[43] In the figure we see that, for both sellers and buyers, these two distributions are very similar to each other. The overall stability in Figure C4 shows that the assumption that the products that firms buy/sell does not change is tenable when analyzing the impact of negative productivity shocks.

3.6.1 How much does the firm-level elasticity of substitution across suppliers matter?

In this section, we assess the importance of the estimated firm-level elasticity of substitution across suppliers for the same product by studying how this elasticity determines the impacts of negative firm-level productivity shocks on aggregate GDP. In this counterfactual, we shock the productivity of firms located in the red zone by 25%. We find that a 25% productivity shock to firms in the red zone reduces GDP by 10.96%. As an empirical benchmark, the state's annual GDP fell by 11.3% in 2020/21. This fall would be 2.017pp less in a model where firms in the same HS-4 product/industry are considered substitutes ($\epsilon = 2$), and 0.75pp more when firms in the same HS-4 product are considered almost Leontief ($\epsilon = 0.001$).

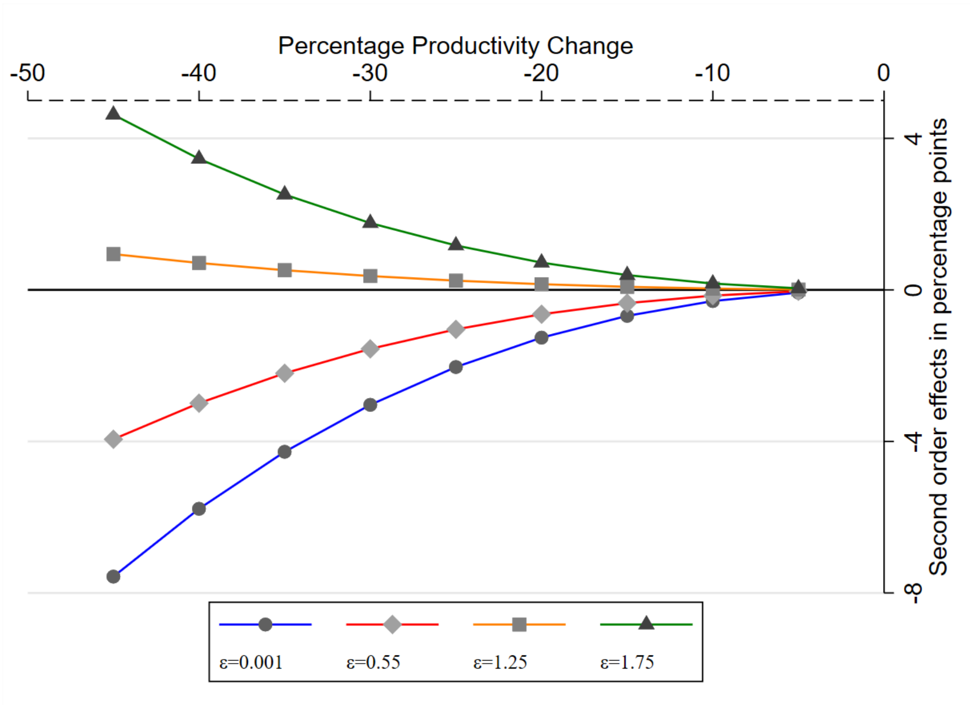
In terms of GDP losses, given that the quarterly GDP of this state was close to 32.5 billion USD in 2020-2021, the additional losses due to firm-level complementarities translate into 655 million USD, which is about 19 USD per capita per quarter, compared to the case when firms are substitutes. To put these numbers into perspective, Baqaee and Farhi (2019) showed that complementarities at the industry level, with an elasticity of substitution 0.001, amplify the effect of a negative 13% shock in the oil-industry on GDP by around 0.61%.

Note that, the differences in GDP that arise from changing values of firm-level elasticities

[43] For the pre-Covid period, t is June 2019-October 2019, and $t-1$ is June 2018-October 2018. For the post-Covid period, t is June 2020-October 2020, and $t-1$ is June 2019-October 2019.

of substitution across suppliers, only changes the second order effects on GDP. Then, how important are these second-order effects that we have estimated? To assess the importance of these second-order effects, in Figure 17 we simulate different levels of negative productivity shocks for 4 different values of the elasticity ϵ and plot the second-order percentage point change in GDP due to these shocks. The top two plots show these differences for high levels of complementarity between suppliers: 0.001 and our estimated elasticity 0.55, respectively. The bottom two plots show the additional change in GDP due to the second order for high levels of substitution across suppliers: 1.25 and 1.75, respectively. Jointly, these plots provide two main lessons. First, for a given negative productivity shock, the second-order effects with the degree of complementarity between suppliers. Second, given the same value of ϵ, the second order effects increase with the magnitude of the productivity shocks. Finally, as suppliers exhibit higher substitutability, the second-order effects actually dampen the negative first-order effects, and more so, for higher values of productivity shocks. That is, unlike the first-order effects which only depend on firm size, complementarities at the firm level non-linearly amplify the effects of negative productivity shocks. This reflects similar amplification patterns that Baqaee and Farhi (2019) documented, but at the industry level.

Figure 17: How important are second-order effects?



Notes: These figures plot the percentage change in productivity to red zone firms on the horizontal axis, and the second-order change in GDP in percentage points on the vertical axis, when the elasticity of substitution across suppliers within the same industry are $\epsilon = 0.001$, $\epsilon = 0.55$, $\epsilon = 1.25$, and $\epsilon = 1.75$, respectively.

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 decades, since Hulten (1978), policy-makers and researchers have emphasized the importance of firm sizes in the propagation of shocks. In the next counterfactual, we investigate how important large firms are versus connected firms in the propagation of shocks.

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

In times of crisis, governments often help small firms stay in business by providing them with subsidies. Nevertheless, governments have limited funds, and it is often not clear for them how to allocate their fixed budget among similarly sized firms. In this counterfactual, we explore the importance of a firm's connectivity in its network when implementing these subsidies. We measure the connectivity of a firm by its value within the Leontief inverse matrix, which measures firms' direct and indirect connections to other firms.[44] Firm size is measured by the size of its domar weight.

Since firm sizes and connectivity are highly correlated with a correlation coefficient of 0.75, we vary the firms' connectivity for a given level of firm size. To implement this, we choose firms that have Domar weights that are equal in size up to 10 decimal places. The first set consists of the most connected firms, the second set is a random draw of firms, and the third set consists of the least connected firms. Since firm sizes are given, the first order effects are the same irrespective of how connected the firms are. In Figure 18, we therefore only plot the second order effects on GDP due to different negative productivity shocks under these three different experiments. In the first scenario, only the most connected firms are affected by negative productivity shocks (blue line). In the second scenario, a random draw of firms is affected (green line). Finally, in the third scenario, only the least connected firms are affected (red line). We perform these experiments under two different elasticities of substitution: an elasticity of substitution amounting to near perfect complementarity ($\epsilon = 0.001$) in the left panel and our estimated complementarity ($\epsilon = 0.55$) in the right panel. All these experiments are conditional on given firm sizes; that is, we vary the connectivity of firms after matching on firm sizes.

These counterfactuals show that the fall in GDP is much larger if the most connected firms are affected compared to the least connected firms, or a random set of firms, for a given firm size. The importance of the most connected firms increases non-linearly with the size of the negative productivity shocks: as the size of the shock increases, it becomes more and more important to give attention to the most connected firms. Our experiment suggests that for our baseline value of elasticity of substitution ($\epsilon = 0.55$) and a negative productivity shock of 45%, if governments save the better-connected firms, given the same firm sizes, compared to randomly targeting firms, the fall in GDP would be about .20 percentage point less, and .31 percentage point less compared to targeting least connected firms.

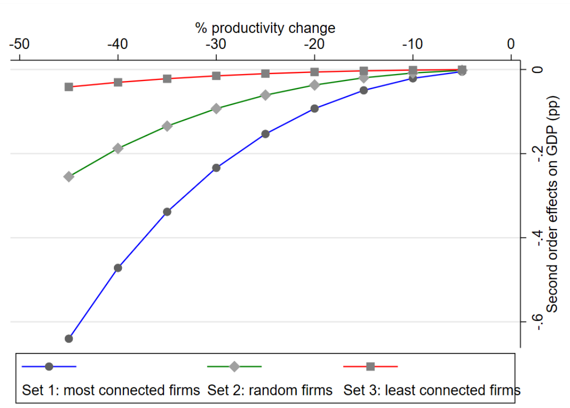
We notice two patterns. First, as the level of the productivity shock reduces, it becomes less important to save the most connected firms. While for a low productivity shock of 5%, the differences in GDP are negligible (.001 and .002), for a productivity shock of 25%, these differences are .05pp and .07 pp compared to saving randomly connected and the least connected firms. Second, the effects of these non-linearities are more pronounced when

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

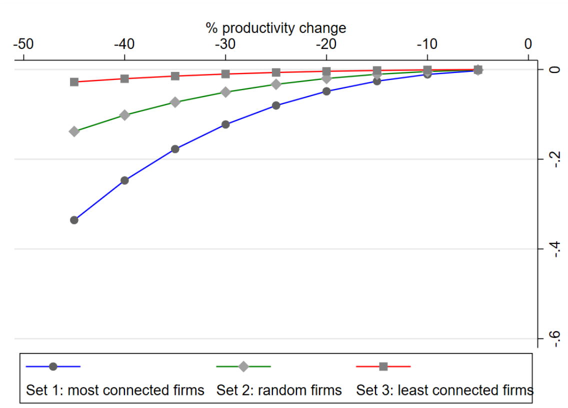
suppliers are highly complementary. For near-perfect complementarity ($\epsilon = 0.001$) and a high negative productivity shock (-45%), the gains from saving the most connected firms compared to saving randomly targeted and least connected firms are .38pp and .60pp, which is almost double the gains if instead suppliers were moderately complementary ($\epsilon = 0.55$). If the goal of policy is to reduce the effects of negative productivity shocks on GDP, for large productivity shocks and low levels of elasticities of substitution, more effective subsidies should target firms that are highly connected.

Figure 18: Second order effects on GDP when firms with same size but different levels of connectivity are affected

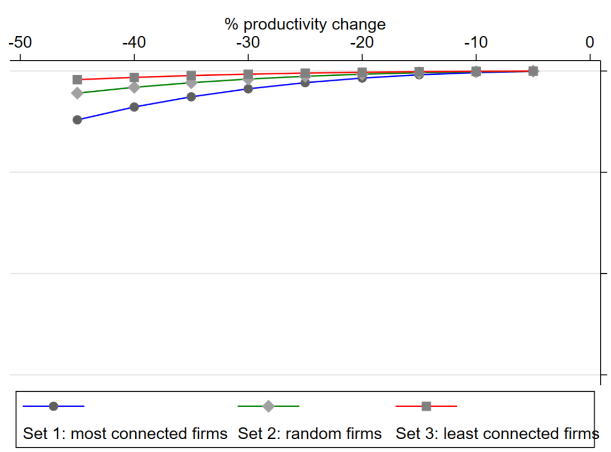
(a) $\epsilon = 0.001$



(b) $\epsilon = 0.55$



(c) $\epsilon = 0.98$



Notes: These figures plot the percentage change in productivity on the horizontal axis, and the second order change in GDP in percentage points on the vertical axis. Sub-figures (a), (b), and (c) plot these effects when the elasticity of substitution across suppliers within the same industry $\epsilon = 0.001$, $\epsilon = 0.55$, and $\epsilon = 0.98$, respectively.

3.6.3 How important is measuring a firm's total connectivity versus direct connectivity?

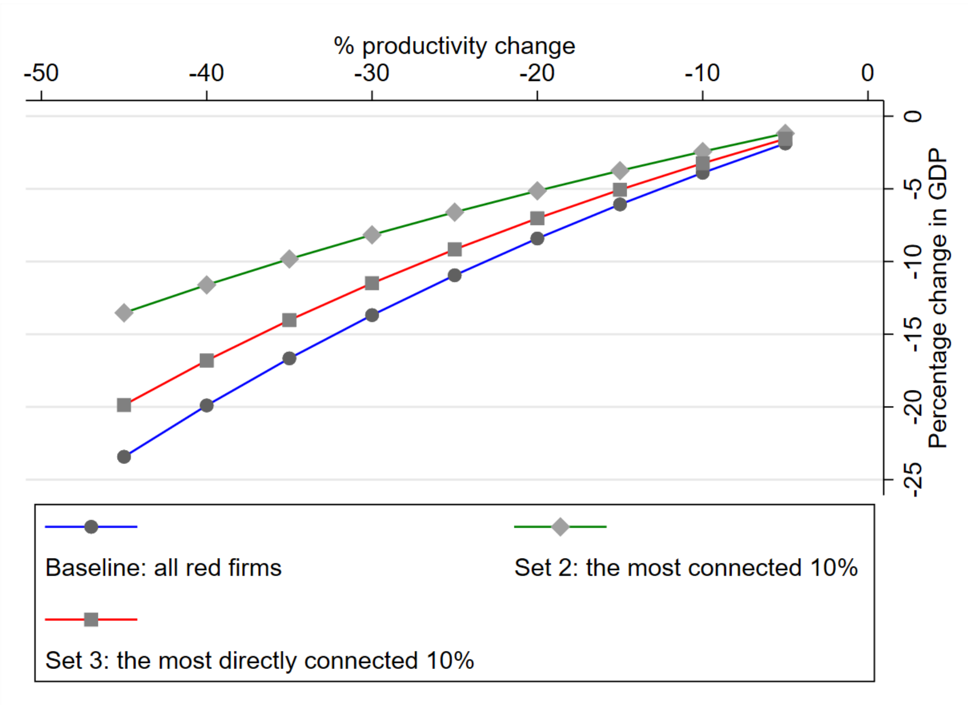
The existing literature has shown that shocks to a firm's suppliers affect the buyer firm and its suppliers (Barrot and Sauvagnat 2018). There is also recent evidence that shocks to a firm can affect its direct as well as other indirect connections (Carvalho et al. 2021). In this counterfactual, we quantify how important it is to take into account a firm's indirect connectivity in understanding how shocks to the firm can affect aggregate GDP. To be precise, a firm's indirect connections measure not only the number of direct buyers of a supplier but also the buyers' buyers and their buyers and so on.[45]

To do this, we conduct two experiments. In the first experiment, the government bails out the most directly connected 10% firms in the red zone, where direct connectivity is measured by the number of buyers a supplier directly supplies (red line in Figure 19). In the second experiment, the government bails the most connected 10\% firms in the red zone, where the total connectivity of a firm is measured by all its direct and indirect connections (green line in Figure 19). Note that, unlike the previous counterfactual, we do not fix firm sizes and vary total connectivity. We are interested in understanding if the government were to bail out just the most directly connected firms as opposed to bailing out the most connected firms irrespective of size, how would that affect aggregate GDP. We report the total effect on the GDP under these two sets of experiments and the baseline results (shock to all red zone firms).

We find that, under our estimated elasticity of $\epsilon = 0.55$ and a negative productivity shock of 25%, the fall in GDP would be 2.56*pp* less if the government were to pick firms on the basis of total connectivity as opposed to direct connectivity. We see that as the level of the negative productivity shock increases, the difference in aggregate GDP between these two sets of experiments rises, emphasizing the importance of measuring a firm's indirect connections as well.

[45] As a reminder, we measure the total connectivity of a firm by its value within the Leontief inverse matrix, which measures firms' direct and indirect connections to other firms. In particular, the $(i, j)^{th}$ entry of the Leontief is a measure of firm i's total reliance on j as a supplier. Summing across all i's yields a measure of the connectivity of each supplier j or its importance in the firm network in terms of connectivity.

Figure 19: Effects on GDP when the largest versus smallest versus all firms are bailed out



Notes: This figure plots the percentage change in productivity of the red zone firms on the horizontal axis and the percentage change in GDP on the vertical axis for our estimated value of $\epsilon = 0.55$. The blue line corresponds to the baseline case when all firms in the red zones are affected. The red line corresponds to the case when the government only bails out the 10% most directly connected firms. The green line corresponds to the case when the government bails out the 10% most connected firms.

3.7 Conclusions

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

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

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

REFERENCES

Abrams, David S, Ufuk Akcigit, and Jillian Grennan (2013). *Patent Value and Citations: Creative Destruction or Strategic Disruption?* Tech. rep. National Bureau of Economic Research.

Acemoglu, Daron et al. (2018). “Innovation, Reallocation, and Growth”. In: *American Economic Review* 108.11, 3450–91.

Akcigit, Ufuk, Douglas Hanley, and Nicolas Serrano-Velarde (2021). “Back to Basics: Basic Research Spillovers, Innovation Policy, and Growth”. In: *The Review of Economic Studies* 88.1, 1–43.

Alfaro-Urena, Alonso, Isabela Manelici, and Jose P Vasquez (2020). “The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages”. In: *Available at SSRN 3376129*.

Allen, Treb and Dave Donaldson (2020). *Persistence and path dependence in the spatial economy*. Tech. rep. National Bureau of Economic Research.

Alviarez, Vanessa et al. (2021). “Two-Sided Market Power in Firm-to-Firm Trade”. In.

Antras, Pol, Teresa C Fort, and Felix Tintelnot (2017). “The margins of global sourcing: Theory and evidence from US firms”. In: *American Economic Review* 107.9, pp. 2514–64.

Arkolakis, Costas, Sun Kyoung Lee, and Michael Peters (2020). *European Immigrants and the United States' Rise to the Technological Frontier*. Tech. rep. mimeo.

Arundel, A., G. van de Paal, and L. Soete (1995). “Innovation Strategies of Europe's Largest Industrial Firms. Results of the PACE Survey for Information Sources, Public Research, Protection of Innovations and Government Programme”. In.

Ash, Elliott et al. (2021). “In-group bias in the Indian judiciary: Evidence from 5 million criminal cases”. In.

Atalay, Enghin (2017). “How Important are Sectoral Shocks?” In: *American Economic Journal: Macroeconomics* 9.4, pp. 254–80.

Atkeson, Andrew and Ariel Tomas Burstein (2010). “Innovation, Firm Dynamics, and International Trade”. In: *Journal of Political Economy* 118.3, 433–484.

Audretsch, David B and Maryann P Feldman (1996). “R&D Spillovers and the Geography of Innovation and Production”. In: *The American Economic Review* 86.3, 630–640.

Bachmann, Ruediger et al. (2022). “What if? The economic effects for Germany of a stop of energy imports from Russia”. In: *ECONtribute Policy Brief* 28, p. 2022.

Bahar, Dany, Prithwiraj Choudhury, and Hillel Rapoport (2020). “Migrant Inventors and the Technological Advantage of Nations”. In: *Research Policy* 49.9, 103947.

Bandyopadhyay, Subhayu, Cletus C Coughlin, and Howard J Wall (2008). “Ethnic networks and US exports”. In: *Review of International Economics* 16.1, pp. 199–213.

Baqaee, David Rezza and Emmanuel Farhi (2019). “The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten’s Theorem”. In: *Econometrica* 87.4, pp. 1155–1203.

Baqaee, David and Emmanuel Farhi (2020). *Nonlinear production networks with an application to the Covid-19 crisis.* Tech. rep. National Bureau of Economic Research.

Barrot, Jean-Noël and Julien Sauvagnat (2018). “Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks”. In: *The Quarterly Journal of Economics* 131.3, pp. 1543–1592.

Behrens, Kristian, Gregory Corcos, and Giordano Mion (2013). “Trade Crisis? What Trade Crisis?” In: *Review of Economics and Statistics* 95.2, pp. 702–709.

Bernard, Andrew B and Andreas Moxnes (2018). “Networks and trade”. In: *Annual Review of Economics* 10, pp. 65–85.

Bernard, Andrew B, Andreas Moxnes, and Yukiko U Saito (2019). “Production networks, geography, and firm performance”. In: *Journal of Political Economy* 127.2, pp. 639–688.

Bernard, Andrew B, Ilke Van Beveren, and Hylke Vandenbussche (2014). “Multi-product exporters and the margins of trade”. In: *The Japanese Economic Review* 65.2, pp. 142–157.

Bernard, Andrew B et al. (2009). “The margins of US trade”. In: *American Economic Review* 99.2, pp. 487–93.

Bernard, Andrew B et al. (2022). “The origins of firm heterogeneity: A production network approach”. In: *Journal of Political Economy* 130.7, pp. 000–000.

Beyer, Robert, Tarun Jain, and Sonalika Sinha (2021). “Lights out? Covid-19 Containment Policies and Economic Activity”. In: *Covid-19 Containment Policies and Economic Activity (May 29, 2021).*

Bhagavatula, Suresh et al. (2018). “Cultural Diversity on Corporate Boards and Firm Outcomes”. In.

Bhattacharyya, Anil (1943). “On a measure of divergence between two statistical populations defined by their probability distributions”. In: *Bulletin of the Calcutta Mathematical Society* 35, pp. 99–109.

Black, Duncan and Vernon Henderson (1999). “A Theory of Urban Growth”. In: *Journal of Political Economy* 107.2, 252–284.

Blind, Knut et al. (2006). “Motives to Patent: Empirical Evidence from Germany”. In: *Research policy* 35.5, 655–672.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen (2013). “Identifying Technology Spillovers and Product Market Rivalry”. In: *Econometrica* 81.4, 1347–1393.

Boehm, Christoph E, Aaron Flaaen, and Nitya Pandalai-Nayar (2019). “Input Linkages and the Transmission of Shocks: Firm-level Evidence from the 2011 Tōhoku Earthquake”. In: *Review of Economics and Statistics* 101.1, pp. 60–75.

Boehm, Johannes and Ezra Oberfield (2020). “Misallocation in the Market for Inputs: Enforcement and the Organization of Production”. In: *The Quarterly Journal of Economics* 135.4, pp. 2007–2058.

Bonadio, Barthélémy et al. (2021). “Global Supply Chains in the Pandemic”. In: *Journal of International Economics*, p. 103534.

Borjas, George J and Kirk B Doran (2012). “The Collapse of the Soviet Union and the Productivity of American Mathematicians”. In: *The Quarterly Journal of Economics* 127.3, 1143–1203.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel (2022). “Quasi-Experimental Shift-share Research Designs”. In: *The Review of Economic Studies* 89.1, 181–213.

Bosetti, Valentina, Cristina Cattaneo, and Elena Verdolini (2015). “Migration of Skilled Workers and Innovation: A European Perspective”. In: *Journal of International Economics* 96.2, 311–322.

Bricongne, Jean-Charles et al. (2012). “Firms and the Global Crisis: French Exports in the Turmoil”. In: *Journal of International Economics* 87.1.

Broda, Christian and David E Weinstein (2006). “Globalization and the Gains from Variety”. In: *The Quarterly Journal of Economics* 121.2, 541–585.

Bryan, Gharad and Melanie Morten (2019). “The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia”. In: *Journal of Political Economy* 127.5, 2229–2268.

Burchardi, Konrad B and Tarek A Hassan (2013). “The Economic Impact of Social Ties: Evidence from German Reunification”. In: *The Quarterly Journal of Economics* 128.3, 1219–1271.

Burchardi, Konrad B et al. (2020). *Immigration, Innovation, and Growth*. Tech. rep. National Bureau of Economic Research.

Cakmakli, Cem et al. (2021). *The Economic Case for Global Vaccinations: An Epidemiological Model with International Production Networks*. Tech. rep. National Bureau of Economic Research.

Carlino, Gerald A, Satyajit Chatterjee, and Robert M Hunt (2007). “Urban Density and the Rate of Invention”. In: *Journal of Urban Economics* 61.3, 389–419.

Carvalho, Vasco M et al. (2021). "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake". In: *The Quarterly Journal of Economics* 136.2.

Chakrabati, Anindya, Kanika Mahajan, and Shekhar Tomar (2021). "Aggregate Shocks, Domestic Trade Collapse and Regional Realignment". In.

Cohen, Wesley M et al. (2002). "R&D Spillovers, Patents and the Incentives to Innovate in Japan and the United States". In: *Research policy* 31.8-9, 1349–1367.

Combes, Pierre-Philippe et al. (2010). "Estimating Agglomeration Economies with History, Geology, and Worker Effects". In: *Agglomeration Economics*. University of Chicago Press, 15–66.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin (2019). "ppmlhdfe: Fast Poisson estimation with high-dimensional fixed effects". In.

Cremers, Katrin et al. (2017). "Patent litigation in Europe". In: *European Journal of Law and Economics* 44.1, pp. 1–44.

Davis, Donald R and Jonathan I Dingel (2019). "A Spatial Knowledge Economy". In: *American Economic Review* 109.1, 153–70.

De Loecker, Jan et al. (2016). "Prices, markups, and trade reform". In: *Econometrica* 84.2, pp. 445–510.

Deloitte (2020). *Germany: Survey of Global Investment and Innovation Incentives*. Tech. rep. Deloitte.

Demir, Banu and Beata Javorcik (2020). "Trade Finance Matters: Evidence from the Covid-19 Crisis". In: *Oxford Review of Economic Policy*.

Demir, Banu et al. (2021). *O-Ring Production Networks*. Tech. rep. National Bureau of Economic Research.

Desmet, Klaus, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg (2018). "The Geography of Development". In: *Journal of Political Economy* 126.3, 903–983.

Desmet, Klaus and Esteban Rossi-Hansberg (2014). "Spatial Development". In: *American Economic Review* 104.4, 1211–43.

Dhyne, Emmanuel, Ayumu Ken Kikkawa, and Glenn Magerman (2022). "Imperfect Competition in Firm-to-Firm Trade". In: *Journal of the European Economic Association* 20.5, pp. 1933–1970.

Dorner, Matthias et al. (2018). *Linked Inventor Biography Data 1980-2014*. Tech. rep. Institut für Arbeitsmarkt-und Berufsforschung (IAB), Nürnberg.

Duguet, Emmanuel and Isabelle Kabla (2000). "Appropriation Strategy and the Motivations to Use the Patent System: an Econometric Analysis at the Firm Level in French Manufacturing". In: *The Economics and Econometrics of Innovation*. Springer, 267–305.

Duranton, Gilles (2007). “Urban Evolutions: The Fast, the Slow, and the Still”. In: *American Economic Review* 97.1, 197–221.

Duranton, Gilles and Diego Puga (2001). “Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products”. In: *American Economic Review* 91.5, 1454–1477.

— (2019). *Urban Growth and its Aggregate Implications*. Tech. rep. National Bureau of Economic Research.

Eaton, Jonathan and Zvi Eckstein (1997). “Cities and Growth: Theory and Evidence from France and Japan”. In: *Regional Science and Urban Economics* 27.4-5, 443–474.

Eaton, Jonathan, Samuel S Kortum, and Francis Kramarz (2022). “Firm-to-Firm Trade: Imports, exports, and the labor market”. In.

Eaton, Jonathan and Samuel Kortum (2002). “Technology, Geography, and Trade”. In: *Econometrica* 70.5, 1741–1779.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz (2011). “An anatomy of international trade: Evidence from French firms”. In: *Econometrica* 79.5, pp. 1453–1498.

Eaton, Jonathan et al. (2016). “Two-sided search in international markets”. In: *2016 Annual Meeting of the Society for Economic Dynamics*.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu (2018). *How Costly are Markups?* Tech. rep. National Bureau of Economic Research.

Fajgelbaum, Pablo D and Edouard Schaal (2020). “Optimal Transport Networks in Spatial Equilibrium”. In: *Econometrica* 88.4, pp. 1411–1452.

Feenstra, Robert C (1994). “New Product Varieties and the Measurement of International Prices”. In: *The American Economic Review*, pp. 157–177.

Feenstra, Robert C, Robert Inklaar, and Marcel P Timmer (2015). “The next generation of the Penn World Table”. In: *American Economic Review* 105.10, pp. 3150–82.

Feldman, Maryann P and Maryellen R Kelley (2006). “The ex ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior”. In: *Research policy* 35.10, pp. 1509–1521.

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

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig (2017). “Cultural proximity and loan outcomes”. In: *American Economic Review* 107.2, pp. 457–92.

Gerschel, Elie, Alejandra Martinez, and Isabelle Mejean (2020). “Propagation of Shocks in Global Value Chains: The Coronavirus Case”. In: *Notes IPP* 53.

Giovanni, Julian di and Andrei A Levchenko (2009). “Trade Openness and Volatility”. In: *The Review of Economics and Statistics* 91.3, pp. 558–585.

Glaeser, Edward L (1999). “Learning in Cities”. In: *Journal of Urban Economics* 46.2, 254–277.

Goolsbee, Austan (1998). *Does Government R&D Policy Mainly Benefit Scientists and Engineers?*

Goraya, Sampreet Singh (2022). “How Does Caste Affect Entrepreneurship?: Birth Vs Worth”. In.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti (2010). “Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings”. In: *Journal of Political Economy* 118.3, 536–598.

Griliches, Zvi (1991). “The Search for R&D spillovers”. In: *National Bureau of Economic Research Working Paper Series* w3768.

Gruber, Jonathan, Simon Johnson, and Enrico Moretti (2022). *Place-Based Productivity and Costs in Science*. Tech. rep. National Bureau of Economic Research.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2009). “Cultural biases in economic exchange?” In: *The Quarterly Journal of Economics* 124.3, pp. 1095–1131.

Hall, Bronwyn H, Adam B Jaffe, and Manuel Trajtenberg (2001). *The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools*.

Hallak, Juan Carlos and Jagadeesh Sivadasan (2013). “Product and Process Productivity: Implications for Quality Choice and Conditional Exporter Premia”. In: *Journal of International Economics* 91.1, 53–67.

Harhoff, Dietmar et al. (1999). “Citation Frequency and the Value of Patented Inventions”. In: *Review of Economics and Statistics* 81.3, 511–515.

Hasanbasri, Ardina (2019). “Explaining the Female Employment Gap across Religions and Castes in India”. In.

Heise, Sebastian et al. (2020). *How did China’s Covid-19 Shutdown affect US Supply Chains?* Tech. rep. Federal Reserve Bank of New York.

Hoisl, Karin et al. (2016). “Social Ties or Patent Quality Signals-Evidence from East German Inventor Migration”. In: *Academy of Management Proceedings*. Vol. 2016. 1. Academy of Management Briarcliff Manor, NY 10510, 13594.

Houthakker, Hendrik S (1955). “The Pareto distribution and the Cobb-Douglas production function in activity analysis”. In: *The Review of Economic Studies* 23.1, pp. 27–31.

Hulten, Charles R (1978). “Growth Accounting with Intermediate Inputs”. In: *The Review of Economic Studies* 45.3, pp. 511–518.

Hummels, David et al. (2014). “The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm data”. In: *American Economic Review* 104.6.

Huneeus, Federico (2018). “Production network dynamics and the propagation of shocks”. In.

Iyer, Lakshmi, Tarun Khanna, and Ashutosh Varshney (2013). “Caste and entrepreneurship in India”. In: *Economic and Political Weekly*, pp. 52–60.

Jaffe, Adam B, Manuel Trajtenberg, and Michael S Fogarty (2000). “Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors”. In: *American Economic Review* 90.2, 215–218.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson (1993). “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations”. In: *the Quarterly Journal of Economics* 108.3, 577–598.

Kantor, Shawn and Alexander Whalley (2019). “Research Proximity and Productivity: Long-term Evidence from Agriculture”. In: *Journal of Political Economy* 127.2, 819–854.

Kelly, Morgan and Anya Hageman (1999). “Marshallian Externalities in Innovation”. In: *Journal of Economic Growth* 4.1, 39–54.

Kerr, William R and Scott Duke Kominers (2015). “Agglomerative Forces and Cluster Shapes”. In: *Review of Economics and Statistics* 97.4, 877–899.

Khanna, Gaurav, Nicolas Morales, and Nitya Pandalai-Nayar (2022). “Supply Chain Resilience: Evidence from Indian Firms”. In.

Kogan, Leonid et al. (2017). “Technological Innovation, Resource Allocation, and Growth”. In: *The Quarterly Journal of Economics* 132.2, 665–712.

Kone, Zovanga L et al. (2018). “Internal Borders and Migration in India”. In: *Journal of Economic Geography* 18.4, pp. 729–759.

Korovkin, Vasily and Alexey Makarin (2020). “Production networks and war”. In: *arXiv preprint arXiv:2011.14756*.

Kortum, Samuel S (1997). “Research, patenting, and technological change”. In: *Econometrica: Journal of the Econometric Society*, 1389–1419.

Kortum, Samuel (1993). “Equilibrium r&d and the patent–r&d ratio: Us evidence”. In: *The American Economic Review* 83.2, 450–457.

Kosfeld, Reinhold and Alexander Werner (2012). “German Labour Markets-New Delineation after the Reforms of German District Boundaries 2007-2011”. In: *RAUMFORSCHUNG UND RAUMORDNUNG-SPATIAL RESEARCH AND PLANNING* 70.1, 49–64.

Krebs, Oliver and Michael Pflüger (2021). “On the road (again): Commuting and local employment elasticities in germany”. In.

Krugman, Paul R (1979). “Increasing returns, monopolistic competition, and international trade”. In: *Journal of international Economics* 9.4, pp. 469–479.

Krugman, Paul (1980). "Scale Economies, Product Differentiation, and the Pattern of Trade". In: *The American Economic Review* 70.5, 950–959.

Kullback, Solomon and Richard A Leibler (1951). "On information and sufficiency". In: *The Annals of Mathematical Statistics* 22.1, pp. 79–86.

Lafrogne-Roussier, R, Julien Martin, and Isabelle Méjean (2021). *Supply Shocks in Supply Chains: Evidence from the Early Lockdown in China*. Tech. rep. Mimeo, CREST.

Lagos, Ricardo (2006). "A model of TFP". In: *The Review of Economic Studies* 73.4, pp. 983–1007.

Lanjouw, Jean O and Mark Schankerman (2004). "Patent quality and research productivity: Measuring innovation with multiple indicators". In: *The economic journal* 114.495, pp. 441–465.

Lim, Kevin (2018). "Endogenous production networks and the business cycle". In.

Macchiavello, Rocco and Ameet Morjaria (2015). "The value of relationships: evidence from a supply shock to Kenyan rose exports". In: *American Economic Review* 105.9, pp. 2911–45.

Marshall, A (1890). "Principles of Economics/Royal Economic Society". In.

Melitz, Marc J (2003). "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity". In: *Econometrica* 71.6, 1695–1725.

Mestieri, Martí, Enrico Berkes, and Ruben Gaetani (2021). "Cities and Technological Waves". In.

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

Monras, Joan (2018). "Economic Shocks and Internal Migration". In.

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg (2018). "Commuting, Migration, and Local Employment Elasticities". In: *American Economic Review* 108.12, 3855–90.

Moretti, Enrico (2021). "The Effect of High-tech Clusters on the Productivity of Top Inventors". In: *American Economic Review* 111.10, 3328–75.

Moser, Petra, Alessandra Voena, and Fabian Waldinger (2014). "German Jewish émigrés and US invention". In: *American Economic Review* 104.10, 3222–55.

Munshi, Kaivan (2014). "Community networks and the process of development". In: *Journal of Economic Perspectives* 28.4, pp. 49–76.

— (2019). "Caste and the Indian economy". In: *Journal of Economic Literature* 57.4, pp. 781–834.

Munshi, Kaivan and Mark Rosenzweig (2016). "Networks and misallocation: Insurance, migration, and the rural-urban wage gap". In: *American Economic Review* 106.1, pp. 46–98.

Nagy, Dávid Krisztián et al. (2016). "City Location and Economic Development". In: *Princeton University, mimeograph.*

Nunn, Nathan (2007). "Relationship-specificity, incomplete contracts, and the pattern of trade". In: *The Quarterly Journal of Economics* 122.2, pp. 569–600.

Oberfield, Ezra (2018). "A theory of input–output architecture". In: *Econometrica* 86.2, pp. 559–589.

Panigrahi, Piyush Paritosh (2021). "Endogenous Spatial Production Networks: Theory, Estimation, and Evidence from Indian Firm-to-Firm Linkages". PhD thesis. University of California, Berkeley.

Papageorgiadis, Nikolaos and Wolfgang Sofka (2020). "Patent enforcement across 51 countries–Patent enforcement index 1998–2017". In: *Journal of World Business* 55.4, p. 101092.

Pellegrina, Heitor S and Sebastian Sotelo (2021). *Migration, Specialization, and Trade: Evidence from Brazil's March to the West.* Tech. rep. National Bureau of Economic Research.

Peri, Giovanni, Kevin Shih, and Chad Sparber (2015). "STEM workers, H-1B visas, and productivity in US cities". In: *Journal of Labor Economics* 33.S1, S225–S255.

Peter, Alessandra and Cian Ruane (2022). "The Aggregate Importance of Intermediate Input Substitutability". In: *Princeton University, memo.*

Peters, Michael (2022). "Market Size and Spatial Growth: Evidence from Germany's Post-war Population Expulsions". In: *Econometrica* 90.5, pp. 2357–2396.

Pitkethly, Robert H (2001). "Intellectual Property Strategy in Japanese and UK Companies: Patent Licensing Decisions and Learning Opportunities". In: *Research Policy* 30.3, 425–442.

Prato, Marta (2021). "The Global Race for Talent: Brain Drain, Knowledge Transfer and Growth". In: *mimeo.*

Ramondo, Natalia, Andrés Rodríguez-Clare, and Milagro Saborío-Rodríguez (2016). "Trade, Domestic Frictions, and Scale Effects". In: *American Economic Review* 106.10, 3159–84.

Rao, Manaswini (2019). "Judges, Lenders, and the Bottom Line: Court-ing Firm Growth in India". In.

Rauch, James E (1996). "Trade and search: Social capital, sogo shosha, and spillovers". In.

— (1999). "Networks versus markets in international trade". In: *Journal of international Economics* 48.1, pp. 7–35.

Rauch, James E and Alessandra Casella (2003). "Overcoming informational barriers to international resource allocation: Prices and ties". In: *The Economic Journal* 113.484, pp. 21–42.

Rauch, James E and Vitor Trindade (2002). "Ethnic Chinese networks in international trade". In: *Review of Economics and Statistics* 84.1, pp. 116–130.

Ravindran, Saravana and Manisha Shah (2020). *Unintended Consequences of Lockdowns: Covid-19 and the Shadow Pandemic*. Tech. rep. National Bureau of Economic Research.

Redding, Stephen J (2016). "Goods trade, factor mobility and welfare". In: *Journal of International Economics* 101, 148–167.

Redding, Stephen J and David E Weinstein (2020). "Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES preferences". In: *The Quarterly Journal of Economics* 135.1.

Richman, Barak D (2006). "How community institutions create economic advantage: Jewish diamond merchants in New York". In: *Law & Social Inquiry* 31.2, pp. 383–420.

Roca, Jorge De La and Diego Puga (2017). "Learning by Working in Big Cities". In: *The Review of Economic Studies* 84.1, 106–142.

Romer, Paul M (2000). "Should the Government Subsidize Supply or Demand in the Market for Scientists and Engineers?" In: *Innovation Policy and the Economy* 1, 221–252.

Schalk, H.J., U.C. Tager, and S Brander (1999). "Wissensverbreitung und Diffusionsdynamik im Spannungsfeld zwischen innovierenden und imitierenden Unternehmen". In: *Ifo-Institut fur Wirtschafts-forschung, Munchen*.

Schoar, Antoinette, Raj Kamal Iyer, and Sandya Kumar (2008). "Importance of Ethnic Networks for Business Transactions of the Small Enterprises". In.

Silverberg, Gerald and Bart Verspagen (2007). "The size distribution of innovations revisited: an application of extreme value statistics to citation and value measures of patent significance". In: *Journal of Econometrics* 139.2, 318–339.

Startz, Meredith (2016). "The value of face-to-face: Search and contracting problems in Nigerian trade". In.

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

Thompson, Peter (2006). "Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations". In: *The Review of Economics and Statistics* 88.2, 383–388.

Trajtenberg, Manuel (1990). "A Penny for your Quotes: Patent Citations and the Value of Innovations". In: *The RAND Journal of Economics*, 172–187.

Trajtenberg, Manuel (2001). “R&D policy in Israel”. In: *Innovation policy in the knowledge-based economy*, pp. 409–454.

WEF (2018). “The global competitiveness report 2018”. In: World Economic Forum.

Wilson, Daniel J (2009). “Beggar Thy Neighbor? The In-state, Out-of-state, and Aggregate Effects of R&D Tax Credits”. In: *The Review of Economics and Statistics* 91.2, 431–436.

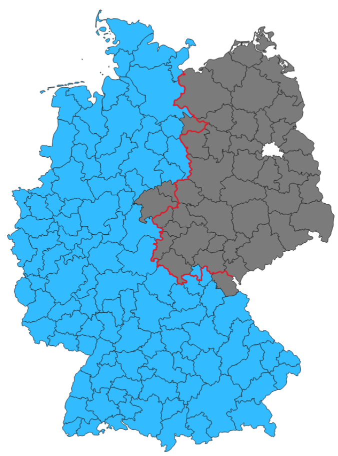
Zhou, Yu (1996). “Inter-firm linkages, ethnic networks, and territorial agglomeration: Chinese computer firms in Los Angeles”. In: *Papers in Regional Science* 75.3, p. 265.

APPENDICES

Appendices to Chapter I

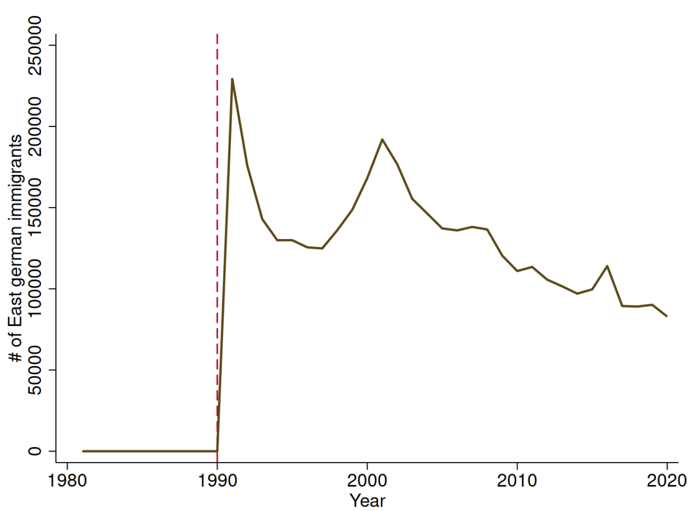
A.1 Additional figures

Figure A1: West Germany



Notes: This figure shows a map of Germany. West Germany is on the left side of the map (blue), and East Germany is on the right side of the map (gray). The red line separating West and East Germany is the *Iron Curtain*, which was lifted on October 1990. The missing area in East Germany is Berlin. The administrative boundaries are labor markets.

Figure A2: The Exodus to the West



Notes: This figure shows the yearly number of East Germans migrating to West Germany. The dashed red line denotes 1990, the date of the Reunification of Germany.

A.2 Additional tables

Table A1: Top 10 West german cities, 2014

	Size		Size
Panel A: Electrical engineering		*Panel D: Mechanical engineering*	
Stuttgart	15.096	Stuttgart	15.865
Munchen	13.776	Munchen	8.140
Regensburg	5.941	Boblingen	5.858
Nurnberg	4.533	Frankfurt	3.577
Erlangen	4.049	Ravensburg	3.318
Karlsruhe	4.005	Erlangen	3.197
Boblingen	2.772	Karlsruhe	3.093
Reutlingen	2.728	Wolsfburg	2.592
Soest	2.552	Dusseldorf	2.540
Frankfurt am Main	2.200	Heilbronn	2.471
Panel B: Instruments		*Panel E: Workers*	
Stuttgart	13.584	Hamburg	6.482
Munchen	8.732	Munchen	5.541
Heidenheim	6.506	Frankfurt	5.369
Erlangen	5.764	Stuttgart	5.070
Boblingen	4.965	Dusseldorf	4.560
Frankfurt	4.109	Koln	3.640
Rottweil	4.052	Essen	3.333
Freiburg	3.424	Hannover	2.541
Regensburg	2.968	Nurnberg	1.932
		Bremen	1.895
Panel C: Chemistry			
Dusseldorf	11.011		
Stuttgart	10.734		
Hamburg	7.202		
Munchen	6.301		
Frankfurt	5.609		
Altotting	2.908		
Essen	2.700		
Koln	2.423		
Reutlingen	2.423		
Erlangen	2.285		

Notes: This table is comprised by five panels. Panels A-D reports the share of inventors working on their corresponding technological area that lives in a given city. Panel E reports the share of workers that lives in a given city. In each panel, I only report the top 10 cities.

Table A2: OLS models, robustness

	Panel A: $\log(1+Z)$					
	(1)	(2)	(3)	(4)	(5)	(6)
EPO	0.117	0.143	0.224	0.184	0.0859	0.173
	(0.0186)	(0.0173)	(0.0135)	(0.0319)	(0.0349)	(0.0679)
EU	0.142	0.193	0.255	0.203	0.103	0.245
	(0.0208)	(0.0162)	(0.0178)	(0.0461)	(0.0463)	(0.0864)
	Panel B: $IHS(Z)$					
	(1)	(2)	(3)	(4)	(5)	(6)
DPMA	0.0847	0.135	0.118	0.130	0.108	0.217
	(0.0326)	(0.0209)	(0.0205)	(0.0475)	(0.0440)	(0.0798)
EPO	0.140	0.171	0.266	0.223	0.102	0.214
	(0.0219)	(0.0204)	(0.0160)	(0.0389)	(0.0431)	(0.0810)
EU	0.142	0.193	0.255	0.203	0.103	0.245
	(0.0208)	(0.0162)	(0.0178)	(0.0461)	(0.0463)	(0.0864)
$\iota_{d,t}$		✓	✓	✓	✓	✓
$\iota_{a,t}$			✓	✓	✓	✓
ι_{da}				✓	✓	✓
ι_{ω}					✓	✓
ι_{i}						✓
N	$177,301$	$177,300$	$177,300$	$177,294$	$162,803$	$84,639$

Notes: In this table I report OLS estimates from Equation (1.1). The table is comprised by two panels. In Panel A, the dependent variable is measured as $\log\left(1+Z_{da,t}^{i\omega}\right)$, where $Z_{da,t}^{i\omega}$ is the number of 5-year forward citations. In Panel B, the dependent variable is measured as $IHS\left(Z_{da,t}^{i\omega}\right)$, where $IHS(\cdot)$ is the inverse hyperbolic sine function. Each panel contains a main set of rows denoted by "DPMA", "EPO", and "EU", which indicate the institution that generated the forward citations. The table is comprised by 6 columns. Rows $3,5,9,11,13$ report the estimate of β, and rows $4,6,10,12,14$ report standard errors clustered at the (d,a) level. Each column corresponds to a different combination of fixed effects, as pointed out by rows $15-19$. Row 20 report the number of observations.

Table A3: OLS models, 5-year periods

	(1)	(2)	(3)	(4)	(5)	(6)
$\log(1+Z)$	0.0291	0.0472	0.0449	0.0707	0.0664	0.0907
	(0.0096)	(0.007)	(0.0073)	(0.0146)	(0.0135)	(0.0215)
$IHS(Z)$	0.0368	0.060	0.0568	0.0902	0.0850	0.116
	(0.0124)	(0.0089)	(0.0094)	(0.0187)	(0.0171)	(0.0273)
$\iota_{d,t}$		✓	✓	✓	✓	✓
$\iota_{a,t}$			✓	✓	✓	✓
ι_{da}				✓	✓	✓
ι_{ω}					✓	✓
ι_{i}						✓
N	$177,301$	$177,300$	$177,300$	$177,294$	$162,803$	$84,639$

Notes: In this table I report OLS estimates from Equation (1.1). Rows $2-3$ report the estimated value of β and its standard errors in parentheses when the dependent variable is measured as $\log\left(1+Z_{da,t}^{i\omega}\right)$, where $Z_{da,t}^{i\omega}$ is the number of 5-year forward citations from the DPMA. Rows $4-5$ report the estimated value of β and its standard errors in parentheses when the dependent variable is measured as $IHS\left(Z_{da,t}^{i\omega}\right)$, where $IHS(\cdot)$ is the inverse hyperbolic sine function. The table is comprised by 6 columns. Each column corresponds to a different combination of fixed effects, as pointed out by rows $6-10$. Standard errors clustered at the (d,a) level. Row 11 reports the number of observations.

Table A4: IV models, robustness

	Panel A: $\Delta \log(1+Z)$		
	(1)	(2)	(3)
EPO	0.164	0.139	0.209
	(0.0422)	(0.0723)	(0.117)
EU	0.210	0.270	0.343
	(0.0436)	(0.0907)	(0.143)
	Panel B: $\Delta IHS(Z)$		
	(1)	(2)	(3)
DPMA	0.215	0.380	0.498
	(0.0514)	(0.122)	(0.184)
EPO	0.182	0.144	0.237
	(0.0494)	(0.0849)	(0.140)
EU	0.235	0.304	0.393
	(0.0588)	(0.104)	(0.168)
$\iota_{d,t}$		✓	✓
$\iota_{a,t}$			✓
$KP-F$	132.1	34.14	28.23
N	$50,778$	$50,776$	$50,776$

Notes: In this table I report IV estimates from Equation (1.2). The table is comprised by two panels. In Panel A, the dependent variable is measured as $\Delta \log\left(1+Z^{i\omega}_{da,t}\right)$, where $Z^{i\omega}_{da,t}$ is the number of 5-year forward citations. In Panel B, the dependent variable is measured as $\Delta IHS\left(Z^{i\omega}_{da,t}\right)$, where $IHS(\cdot)$ is the inverse hyperbolic sine function. Each panel contains a main set of rows denoted by "DPMA", "EPO", and "EU", which indicate the institution that generated the forward citations. The table is comprised by 3 columns. Rows $3, 5, 9, 11, 13$ report the estimate of β, and rows $4, 6, 10, 12, 14$ report standard errors clustered at the (d, a) level. Each column corresponds to a different combination of fixed effects, as pointed out by rows $15-16$. Row 19 shows the first stage Kleibergen-Paap F-statistic (KP-F), and row 20 reports the number of observations.

Table A5: IV models, 5-year periods

	(1)	(2)	(3)
$\Delta \log (1+Z)$	0.0367	0.0865	0.0849
	(0.0232)	(0.0331)	(0.0428)
$\Delta IHS(Z)$	0.0464	0.109	0.104
	(0.0295)	(0.0420)	(0.0543)
$\iota_{d,t}$		✓	✓
$\iota_{a,t}$			✓
$KP-F$	85.96	26.64	38.15
N	$100,234$	$100,228$	$100,228$

Notes: In this table I report IV estimates from Equation (1.2). Rows $2-3$ report the estimated value of β and its standard errors in parentheses when the dependent variable is measured as $\Delta \log \left(1+Z_{da,t}^{i\omega}\right)$, where $Z_{da,t}^{i\omega}$ is the number of 5-year forward citations from the DPMA. Rows $4-5$ report the estimated value of β and its standard errors in parentheses when the dependent variable is measured as $\Delta IHS\left(Z_{da,t}^{i\omega}\right)$, where $IHS(\cdot)$ is the inverse hyperbolic sine function. The table is comprised by 3 columns. Each column corresponds to a different combination of fixed effects, as pointed out by rows $6-7$. Standard errors clustered at the (d,a) level. Row 8 shows the first stage Kleibergen-Paap F-statistic (KP-F), and row 9 reports the number of observations.

Table A6: Estimation of decreasing returns to R&D, 5-year periods

	(1)	(2)	(3)
$\log\left(R_{o,t}^{\omega}\right)$	0.661	0.644	0.568
	(0.008)	(0.0081)	(0.0098)
$\iota_{o,t}$		✓	✓
ι_{ω}			✓
N	$95,699$	$95,699$	$68,381$
R^2	0.683	0.744	0.854

Notes: In this table I report estimates for decreasing returns to R&D from Equation (1.35). The dependent variable log $\left(n_{o,t}^{\omega}\right)$ is the number of firm's inventors that filed a patent. Time are 5-year periods. Each column is an specification with different combinations of fixed effects. The fixed effects included in each specification are determined by rows $4-5$. Row 2 contains the estimates for ζ, and row 3 contain standard errors, which are clustered at the o,t level. Rows $6-7$ contain the number of observations and goodness of fit in each specification, respectively.

Table A7: Estimation of migration costs, 5-year periods

	$n = R$		$n = L$	
	OLS	PPML	OLS	PPML
$\log\left(dist_{od}\right)$	-1.020	-1.381	-1.505	-1.380
	(0.010)	(0.017)	(0.025)	(0.015)
ρ_1^n	0.481	0.651	0.709	0.650
R^2	0.826	$\cdot$	0.835	$\cdot$
N	$17,283$	$54,080$	$43,835$	$54,080$

Notes: In this table I report migration cost elasticities from Equation (1.36) under 5-year periods. Columns $2-3$ are the regressions for inventors, where column 2 are OLS estimates, and column 3 are PPML estimates. Columns $4-5$ are the regressions for workers, where column 4 are OLS estimates, and column 5 are PPML estimates. For OLS estimates, the dependent variable is measured as $\log\left(\eta_{od,t}^n\right)$ is the log of the share of inventors or workers from o that moved to d during a given period, where I consider 5-year periods. Row 3 is the estimate associated to $\log\left(dist_{od}\right)$, where $dist_{od}$ is the Euclidean distance in miles from o to d. Row 4 are standard errors two-way clustered at the o, t and d, t level. Row 5 is the implied migration elasticity from the estimates from row 3 given $\kappa = 2.12$. Rows $6-7$ contain the goodness of fit and number of observations in each specification, respectively.

A.3 Linked inventor biography data (INV-BIO)

The INV-BIO is comprised by approximately $150,000$ inventors in Germany with high–frequency information on their employment spells and patenting activities between 1980 and 2014. All inventors recorded in the INV-BIO data filed at least one patent with the European Patent Office (EPO) between 1999 and 2011 and were disambiguated using a combination of record linkage and machine learning methods.[46] The INV-BIO dataset is comprised by three modules: (i) inventor module, (ii) establishment module, and (iii) patent module. I now describe details of each module.

A.3.1 Module on inventors.

The module on inventors is reported at the employment spell level. I now explain how I collapse the data at the inventor and period level. For a given inventor and year, consider the set of the inventor's spells. Then, for a given spell, the data contains information on the establishment an inventor works for, inventor's daily wage, 1-digit occupation code, whether inventor's job is part time, and the inventor's residence location. Since it is possible that an inventor reports multiple jobs within a year, an inventor's job is the one with the longest tenure. Whenever a tie happens, an inventor's job is the one with the highest daily wage. If a tie still remains, an inventor's job is chosen randomly. Part-time jobs are excluded.

[46]For more details, see Dorner et al. (2018)

Finally, when collapsing the data at the inventor and period level, the last year within a period defines inventor characteristics.

A.3.2 Module on establishments.

The module on establishments is reported at the establishment and year level. I now explain how I collapse the data at the establishment and period level. The data contains a 1-digit 2008 time-consistent NACE code, the year an establishment is registered in the German administrative records for the first time, the year an establishment stops being registered in the German administrative records, and establishment location. Then, a panel of establishments is constructed based on the years the establishments were first and last registered. If the first year an establishment is registered in the data is before 1980, data on that establishment begins on 1980. If the last year an establishment is registered in the data is after 2014, data on that establishment ends on 2014. It is possible that an establishment is not registered in a given year because of lack of patenting activity by its inventors. Whenever that happens, an establishment is considered to still exists during those years, such that their industry and location are the same from the previous year. Finally, when collapsing the data at the inventor and period level, the last year within a period defines establishment characteristics.

A.3.3 Module on patents.

The module on patents is reported at the patent and inventor level. I now explain how I collapse the data at the inventor and period level. The data contains patent characteristics such as the date when the patent was filed for the first time, 2-10 forward year citations from the German Patent and Trade Mark Office (DPMA), the European Patent Office (EPO), and the United States Patent and Trademark Office (US); the mean distance between the inventors that filed the patent, 1-digit technological area, and originality and generality indices. For each patent, the earliest filing date determines the year when the patent was generated. Then, the data is collapsed at the inventor and year, such that the data reports the number of forward citations and number of filed patents during a given year. Finally, when collapsing at the inventor and period, the number of forward citations and number of filed patents are added up.

A.4 Microfoundations

In this section I provide details on the microfoundations of the model.

A.4.1 Generation of ideas

Consider an inventor i working for firm ω in location o. Consider $Z_o^{i\omega,j}$ to be the productivity of an idea j that inventor i generated. As in Kortum (1997), innovation is the process where an inventor generates T_o ideas and selects the one with the highest productivity, such that

$$Z_o^{i\omega} = \max_{j=1,\ldots,T_o} Z_o^{i\omega,j}.$$

Then, the conditional probability distribution of inventor i's best idea is

$$\begin{aligned}
G\left(z \mid T_o\right) &= \Pr\left\{Z_o^{i\omega,j} \leq z \mid T_o\right\}, \\
&= \Pr\left\{Z_o^{i\omega,1} \leq z, \ldots, Z_o^{i\omega,T_o} \leq z \mid T_o\right\}, \\
&= \Pr\left\{Z_o^{i\omega,1} \leq z\right\} \times \cdots \times \Pr\left\{Z_o^{i\omega,T_o} \leq z\right\}, \\
&= \underbrace{F\left(z\right) \times \cdots \times F\left(z\right)}_{T_o\ times}, \\
&= F\left(z\right)^{T_o},
\end{aligned}$$

where $F\left(z\right)$ is the cumulative probability that an idea drawn by inventor i is below productivity z. Since T_o is the discrete number of ideas drawn by and inventor, I assume that T_o follows a Poisson distribution, such that $\Pr\left\{T_o = n\right\} = \frac{\lambda_o^n \exp(-\lambda_o)}{n!}$, where n is the number of drawn ideas, and λ_o is the expected number of drawn ideas. Additionally, I assume that ideas are drawn from a Pareto distribution, such that $F\left(z\right) = 1 - z^{-\alpha}$, where $\alpha > 1$ is a shape parameter. Then, the unconditional distribution of the productivity of inventor i's best idea is

$$\begin{aligned}
G\left(z\right) &= \Pr\left\{Z_o^{i\omega} \leq z\right\}, \\
&= \sum_{n=0}^{\infty} \left[\frac{\lambda_o^n \exp\left(-\lambda_o\right)}{n!}\right] \left[F\left(z\right)^n\right], \\
&= \exp\left(-\lambda_o\right) \left[\sum_{n=0}^{\infty} \frac{\left(\lambda_o F\left(z\right)\right)^n}{n!}\right], \\
&= \exp\left(-\lambda_o\right) \exp\left(\lambda_o F\left(z\right)\right), \\
&= \exp\left(-\lambda_o \left(1 - F\left(z\right)\right)\right), \\
&= \exp\left(-\lambda_o \left(1 - \left(1 - z^{-\alpha}\right)\right)\right), \\
&= \exp\left(-\lambda_o z^{-\alpha}\right).
\end{aligned}$$

That is, $Z_o^{i\omega}$ is drawn from a Frechet distribution with shape parameter α and scale $\lambda_o^{\frac{1}{\alpha}}$.

A.4.2 Microfoundations: quality of intermediate inputs

Microfoundation 1: necessary tasks. Consider that a unit of the intermediate input ω is produced at a level of quality determined by a blueprint. The firm that produces its unique input ω owns the blueprint. A blueprint is defined as a continuum of tasks $\mathcal{T} \equiv [0,1]$ that are necessary to produce the input at a given quality. Then, the quality of the unit of an intermediate input is

$$A_o^\omega = \exp\left(\int_{\mathcal{T}} \log\left(A_o^{\omega,\tau}\right) d\tau\right),$$

where $A_o^{\omega,\tau}$ is the quality of task $\tau \in \mathcal{T}$ within ω's blueprint. The firm hires a mass of inventors R_o^ω who generate $n_o^\omega \leq R_o^\omega$ ideas that determine the quality of each task within the ω's blueprint. Ideas are heterogeneous in productivity and each idea improves the quality of all tasks within the blueprint, such that the quality of each task is

$$A_o^{\omega,\tau} = z^\tau n_o^\omega,$$

where z^τ is the productivity of each idea generated by firms' inventors. Plugging this into the expression for A_o^ω yields

$$\begin{aligned}
A_o^\omega &= \exp\left(\int_{\mathcal{T}} \log\left(A_o^{\omega,\tau}\right) dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau n_o^\omega\right) dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \left[\log\left(z^\tau\right) + \log\left(n_o^\omega\right)\right] dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau + \int_{\mathcal{T}} \log\left(n_o^\omega\right) dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau\right) \exp\left(\int_{\mathcal{T}} \log\left(n_o^\omega\right) dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau\right) \exp\left(\log\left(n_o^\omega\right) \int_{\mathcal{T}} dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau\right) \exp\left(\log\left(n_o^\omega\right)\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau\right) n_o^\omega.
\end{aligned}$$

Since z^{τ} are draws from a Frechet distribution as in Equation (1.13), then $\log(z^{\tau})$ are draws from a Gumbel distribution with location parameter $\log\left(\lambda_o^{\frac{1}{\alpha}}\right)$ and scale parameter $\frac{1}{\alpha}$. Then,

$$
\begin{aligned}
A_o^{\omega} &= \exp\left(\int_{\mathcal{T}} \log(z^{\tau})\, d\tau\right) n_o^{\omega}, \\
&= \exp\left(\int_0^{\infty} \log(z)\, dG_o(z)\right) n_o^{\omega}, \\
&= \exp\left(\log\left(\lambda_o^{\frac{1}{\alpha}}\right) + \frac{\overline{\gamma}}{\alpha}\right) n_o^{\omega}, \\
&= \exp\left(\log\left(\lambda_o^{\frac{1}{\alpha}}\right)\right) \exp\left(\frac{\overline{\gamma}}{\alpha}\right) n_o^{\omega}, \\
&= \psi \lambda_o^{\frac{1}{\alpha}} n_o^{\omega}.
\end{aligned}
$$

where $\psi \equiv \exp\left(\frac{\overline{\gamma}}{\alpha}\right)$ is a constant, and $\overline{\gamma}$ is Euler's constant.

Microfoundation 2: linear innovation. Consider that a unit of the intermediate input ω is produced at a level of quality determined by a blueprint. The firm that produces its unique input ω owns the blueprint. A blueprint is defined as a the average quality of all the ideas generated by firms' inventors. Consider a firm ω that hires a mass of inventors R_o^{ω}. The task of each inventor is to come up with an idea that will be incorporated into the firm's blueprint. Inventors show up for work, they form an arbitrary line, the first inventor receives the blueprint, implements his idea into the blueprint and passes it over to the next inventor, and so on. At the end of the line, $n_o^{\omega} \leq R_o^{\omega}$ ideas have been implemented into the blueprint since some inventors are not able to generate an idea due decreasing returns to R&D (e.g. duplication effects). Ideas are heterogeneous in productivity since they are drawn from a Frechet distribution as in (1.13). Then, the quality of intermediate input ω is

$$
\begin{aligned}
A_o^{\omega} &= \int_o^{n_o^{\omega}} z^i di, \\
&= n_o^{\omega} \int_o^{\infty} z dG(z), \\
&= n_o^{\omega} \left[\Gamma\left(1 - \frac{1}{\alpha}\right) \lambda_o^{\frac{1}{\alpha}}\right], \\
&= \psi \lambda_o^{\frac{1}{\alpha}} n_o^{\omega},
\end{aligned}
$$

where $\psi \equiv \Gamma\left(1 - \frac{1}{\alpha}\right)$ is a constant, and $\Gamma(\cdot)$ is the Gamma function.

A.5 Derivations

Final good firms. In each location d, a representative firm produces a final good by aggregating intermediates from all locations. The production function of the final good is

$$Q_d = \left(\sum_o \int_{\omega \in \Omega_{od}} A_o^{\omega \frac{1}{\sigma}} Q_{od}^{\omega \frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \tag{A1}$$

where Ω_{od} is the set of intermediate input firms in o selling to d, Q_d is the production of the final good, Q_{od}^{ω} is the quantity of intermediate input ω, A_o^{ω} is input quality, and $\sigma > 1$ is the CES across intermediate inputs. The final good producer maximizes profits:

$$\max_{\{Q_{od}^{\omega}\}} \; P_d Q_d - \sum_o \int_{\omega \in \Omega_{od}} P_{od}^{\omega} Q_{od}^{\omega} s.t.$$
$$Q_d = \left(\sum_o \int_{\omega \in \Omega_{od}} A_o^{\omega \frac{1}{\sigma}} Q_{od}^{\omega \frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}.$$

The first order condition of buying an intermediate input ω from o is

$$\begin{aligned}
[Q_{od}^{\omega}] : & P_d \left(\frac{\sigma}{\sigma-1} \right) (\ldots_d)^{\frac{\sigma}{\sigma-1}-1} A_o^{\omega \frac{1}{\sigma}} \left(\frac{\sigma-1}{\sigma} \right) Q_{od}^{\omega \frac{\sigma-1}{\sigma}-1} = P_{od}^{\omega}, \\
P_{od}^{\omega} &= P_d (\ldots_d)^{\frac{\sigma-\sigma+1}{\sigma-1}} A_o^{\omega \frac{1}{\sigma}} Q_{od}^{\omega \frac{\sigma-1-\sigma}{\sigma}}, \\
&= P_d (\ldots_d)^{\frac{1}{\sigma-1}} A_o^{\omega \frac{1}{\sigma}} Q_{od}^{\omega -\frac{1}{\sigma}},
\end{aligned}$$

where $(\ldots_d)$ is a composite of terms in d. Now, consider the first order condition of buying an intermediate input ω from o':

$$P_{od}^{\omega} = P_d (\ldots_d)^{\frac{1}{\sigma-1}} A_o^{\omega \frac{1}{\sigma}} Q_{o'd}^{\omega -\frac{1}{\sigma}}.$$

Divide both order conditions:

$$
\begin{aligned}
\frac{P_{od}^{\omega}}{P_{o'd}^{\omega}} &= \frac{P_d\left(\ldots_d\right)^{\frac{1}{\sigma-1}} A_o^{\omega\frac{1}{\sigma}} Q_{od}^{\omega-\frac{1}{\sigma}}}{P_d\left(\ldots_d\right)^{\frac{1}{\sigma-1}} A_o^{\omega\frac{1}{\sigma}} Q_{o'd}^{\omega-\frac{1}{\sigma}}}, \\
&= \frac{A_o^{\omega\frac{1}{\sigma}} Q_{od}^{\omega-\frac{1}{\sigma}}}{A_{o'}^{\omega\frac{1}{\sigma}} Q_{o'd}^{\omega-\frac{1}{\sigma}}}, \\
&= \frac{A_o^{\omega\frac{1}{\sigma}} Q_{o'd}^{\omega\frac{1}{\sigma}}}{A_{o'}^{\omega\frac{1}{\sigma}} Q_{od}^{\omega\frac{1}{\sigma}}}, \\
\frac{P_{od}^{\omega\sigma-1}}{P_{o'd}^{\omega\sigma-1}} &= \frac{A_o^{\omega\frac{\sigma-1}{\sigma}} Q_{o'd}^{\omega\frac{\sigma-1}{\sigma}}}{A_{o'}^{\omega\frac{\sigma-1}{\sigma}} Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}, \\
Q_{o'd}^{\omega\frac{\sigma-1}{\sigma}} &= \frac{A_{o'}^{\omega\frac{\sigma-1}{\sigma}} Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_o^{\omega\frac{\sigma-1}{\sigma}}} \frac{P_{od}^{\omega\sigma-1}}{P_{o'd}^{\omega\sigma-1}}.
\end{aligned}
$$

Plug this expression in the production function of the final good producer:

$$
\begin{aligned}
Q_d &= \left(\sum_{o'} \int_{\omega\in\Omega_{o'd}} A_{o'}^{\omega\frac{1}{\sigma}} Q_{o'd}^{\omega\frac{\sigma-1}{\sigma}} d\omega\right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\sum_{o'} \int_{\omega\in\Omega_{o'd}} A_{o'}^{\omega\frac{1}{\sigma}} \frac{A_{o'}^{\omega\frac{\sigma-1}{\sigma}} Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_o^{\omega\frac{\sigma-1}{\sigma}}} \frac{P_{od}^{\omega\sigma-1}}{P_{o'd}^{\omega\sigma-1}} d\omega\right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\frac{Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_o^{\omega\frac{\sigma-1}{\sigma}}} P_{od}^{\omega\sigma-1} \sum_{o'} \int_{\omega\in\Omega_{o'd}} A_{o'}^{\omega\frac{1}{\sigma}} A_{o'}^{\omega\frac{\sigma-1}{\sigma}} P_{o'd}^{\omega 1-\sigma} d\omega\right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\frac{Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_o^{\omega\frac{\sigma-1}{\sigma}}} P_{od}^{\omega\sigma-1} \sum_{o'} \int_{\omega\in\Omega_{o'd}} A_{o'}^{\omega} P_{o'd}^{\omega 1-\sigma} d\omega\right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\frac{Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_o^{\omega\frac{\sigma-1}{\sigma}}} P_{od}^{\omega\sigma-1} \sum_{o'} P_{od,t}^{1-\sigma}\right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\frac{Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_o^{\omega\frac{\sigma-1}{\sigma}}} P_{od}^{\omega\sigma-1} P_{d,t}^{1-\sigma}\right)^{\frac{\sigma}{\sigma-1}},
\end{aligned}
$$

where $P_{d,t}^{1-\sigma} = \left(\sum_o P_{od,t}^{1-\sigma}\right)$, and $P_{od,t}^{1-\sigma} = \left(\int_{\omega \in \Omega_{od,t}} A_o^\omega P_{od}^{\omega^{1-\sigma}} d\omega\right)$ are CES price indices. Then, rearrange this expression to obtain the demand of intermediate inputs from o:

$$
\begin{aligned}
Q_d &= \left(\frac{Q_{od}^{\omega^{\frac{\sigma-1}{\sigma}}}}{A_o^{\omega^{\frac{\sigma-1}{\sigma}}}} P_{od}^{\omega^{\sigma-1}} P_{d,t}^{1-\sigma} \right)^{\frac{\sigma}{\sigma-1}}, \\
&= \frac{Q_{od}^\omega}{A_o^\omega} P_{od}^{\omega^\sigma} P_{d,t}^{-\sigma}, \\
Q_{od}^\omega &= A_o^\omega P_{od}^{\omega^{-\sigma}} P_d^\sigma Q_d, \\
&= A_o^\omega P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} \left(P_d Q_d\right), \\
&= A_o^\omega P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d,
\end{aligned}
$$

where $X_d = P_d Q_d$ is total expenditure of the final good in d.

Intermediate input firms. The intermediate input firm in o maximizes profits by selling its inputs to all locations subject to the demand from every location and its cost structure:

$$
\begin{aligned}
\max_{\left\{P_{od}^\omega, Q_{od}^\omega, L_{od}^\omega\right\}} \pi_o^\omega &= \sum_d \pi_{od}^\omega, \\
&s.t. \\
\pi_{od}^\omega &= P_{od}^\omega Q_{od}^\omega - \tau_{od} w_o^L L_{od}^\omega, \\
L_{od}^\omega &= Q_{od}^\omega, \\
Q_{od}^\omega &= A_o^\omega P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d.
\end{aligned}
$$

Introduce the constraints into the profit function:

$$
\begin{aligned}
\pi_o^\omega &= \sum_d \pi_{od}^\omega, \\
&= \sum_d \left(P_{od}^\omega Q_{od}^\omega - \tau_{od} w_o^L L_{od}^\omega \right), \\
&= \sum_d \left(P_{od}^\omega Q_{od}^\omega - \tau_{od} w_o^L Q_{od}^\omega \right), \\
&= \sum_d \left(P_{od}^\omega A_o^\omega P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d \right) \\
&- \sum_d \left(\tau_{od} w_o^L A_o^\omega P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d \right), \\
&= \sum_d \left(A_o^\omega P_{od}^{\omega^{1-\sigma}} P_d^{\sigma-1} X_d \right) \\
&- \sum_d \left(\tau_{od} w_o^L A_o^\omega P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d \right).
\end{aligned}
$$

The first order condition is

$$
\begin{aligned}
[P_{od}^\omega] : (1-\sigma)\, A_o^\omega P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d - (-\sigma)\, \tau_{od} w_o^L A_o^\omega P_{od}^{\omega^{-\sigma-1}} P_d^{\sigma-1} X_d &= 0, \\
0 &= (1-\sigma)\, P_{od}^{\omega^{-\sigma}} + \sigma \tau_{od} w_o^L P_{od}^{\omega^{-\sigma-1}}, \\
(\sigma - 1) &= \sigma \tau_{od} w_o^L P_{od}^{\omega^{-1}}, \\
P_{od}^\omega &= \left(\frac{\sigma}{\sigma - 1} \right) \tau_{od} w_o^L, \\
&= \overline{m} \tau_{od} w_o^L,
\end{aligned}
$$

where $\overline{m} \equiv \frac{\sigma}{\sigma-1}$ is the CES constant markup over marginal costs.

Total profits. Introducing the markup pricing Equation (1.9) in the profit function (1.8) yields

$$
\begin{aligned}
\pi_o^\omega &= \sum_d \pi_{od}^\omega, \\
&= \sum_d P_{od}^\omega Q_{od}^\omega - \tau_{od} w_o^L L_{od}^\omega, \\
&= \sum_d P_{od}^\omega Q_{od}^\omega - \tau_{od} w_o^L Q_{od}^\omega, \\
&= \sum_d \left(P_{od}^\omega - \tau_{od} w_o^L\right) Q_{od}^\omega, \\
&= \sum_d \left(\overline{m}\tau_{od} w_o^L - \tau_{od} w_o^L\right) Q_{od}^\omega, \\
&= \sum_d \left(\overline{m} - 1\right) \tau_{od} w_o^L Q_{od}^\omega, \\
&= \sum_d \left(\overline{m} - 1\right) \tau_{od} w_o^L A_o^\omega P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d, \\
&= \sum_d \left(\overline{m} - 1\right) \tau_{od} w_o^L A_o^\omega \left(\overline{m}\tau_{od} w_o^L\right)^{-\sigma} P_d^{\sigma-1} X_d, \\
&= \sum_d \left(\overline{m} - 1\right) \overline{m}^{-1} \left(\overline{m}\tau_{od} w_o^L\right) A_o^\omega \left(\overline{m}\tau_{od} w_o^L\right)^{-\sigma} P_d^{\sigma-1} X_d, \\
&= A_o^\omega \sum_d \left(\overline{m} - 1\right) \overline{m}^{-1} \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma} P_d^{\sigma-1} X_d, \\
&= \left(\frac{\overline{m} - 1}{\overline{m}}\right) A_o^\omega \sum_d \left(P_{od}^\omega\right)^{1-\sigma} P_d^{\sigma-1} X_d, \\
&= \left(\frac{\frac{\sigma}{\sigma-1} - 1}{\frac{\sigma}{\sigma-1}}\right) A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d, \\
&= \left(\frac{\frac{1}{\sigma-1}}{\frac{\sigma}{\sigma-1}}\right) A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d, \\
&= \frac{1}{\sigma} A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d.
\end{aligned}
$$

Research and Development (R&D). Firm ω maximizes total profits after R&D expenditure subject to its profits before R&D and the quality of its intermediate:

$$\begin{aligned}
\max_{\{R_o^\omega\}} \overline{\pi}_o^\omega &= \pi_o^\omega - w_o^R R_o^\omega \\
& s.t. \\
\pi_o^\omega &= \frac{1}{\sigma} A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d, \\
A_o^\omega &= \psi \mathcal{A}_o R_o^{\widetilde{\gamma}} R_o^{\omega\zeta}.
\end{aligned}$$

Rewrite profits:

$$\begin{aligned}
\overline{\pi}_o^\omega &= \pi_o^\omega - w_o^R R_o^\omega, \\
&= \frac{1}{\sigma} A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d - w_o^R R_o^\omega, \\
&= \frac{1}{\sigma} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right) n_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d - w_o^R R_o^\omega, \\
&= \frac{1}{\sigma} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right) R_o^{\omega\zeta} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d - w_o^R R_o^\omega.
\end{aligned}$$

The first order condition is

$$\begin{aligned}
[R_o^\omega] : w_o^R &= \frac{\zeta}{\sigma} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right) R_o^{\omega\zeta-1} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d, \\
R_o^{\omega 1-\zeta} &= \frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d, \\
R_o^\omega &= \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}.
\end{aligned}$$

The number of implemented ideas is

$$\begin{aligned}
n_o^\omega &= R_o^{\omega\zeta}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{\zeta}{1-\zeta}}.
\end{aligned}$$

The quality of the intermediate is

$$
\begin{aligned}
A_o^\omega &= \psi \mathcal{A}_o R_o^{\widetilde{\gamma}} n_o^\omega, \\
&= \psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{\frac{1-\zeta}{\zeta}} \sum_d \left(\frac{P_{od}^\omega}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{1+\frac{1-\zeta}{\zeta}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{\frac{\zeta+1-\zeta}{\zeta}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{\frac{1}{\zeta}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{\frac{1}{\zeta}}}{w_o^R} \sum_d \left(\frac{\overline{m} \tau_{od} w_o^L}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{\frac{1}{\zeta}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{\zeta}{1-\zeta}}.
\end{aligned}
$$

Then, total profits after R&D is

$$
\begin{aligned}
\overline{\pi}_o^\omega &= \frac{1}{\sigma} A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d - w_o^R R_o^\omega, \\
&= \frac{1}{\sigma} \left(\frac{\zeta}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{\zeta}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \\
&- w_o^R \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}}, \\
&= \frac{1}{\sigma} \left(\frac{\zeta}{\sigma} w_o^{R^{-1}} \right)^{\frac{\zeta}{1-\zeta}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{1+\frac{\zeta}{1-\zeta}} \\
&- w_o^R \left(\frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \right)^{\frac{1}{1-\zeta}} \left(\frac{\zeta}{\sigma} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}}, \\
&= \frac{1}{\sigma} \left(\frac{\zeta}{\sigma} \right)^{\frac{\zeta}{1-\zeta}} \left(w_o^R\right)^{\frac{\zeta}{\zeta-1}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}} \\
&- w_o^R \left(w_o^R\right)^{-\frac{1}{1-\zeta}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\frac{\zeta}{\sigma} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}}, \\
&= \frac{\zeta^{\frac{\zeta}{1-\zeta}}}{\sigma^{1+\frac{\zeta}{1-\zeta}}} \left(w_o^R\right)^{\frac{\zeta}{\zeta-1}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}} \\
&- \left(\frac{\zeta}{\sigma} \right)^{\frac{1}{1-\zeta}} \left(w_o^R\right)^{1-\frac{1}{1-\zeta}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}}, \\
&= \frac{\zeta^{\frac{\zeta}{1-\zeta}}}{\sigma^{\frac{1}{1-\zeta}}} \left(w_o^R\right)^{\frac{\zeta}{\zeta-1}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}} \\
&- \frac{\zeta^{\frac{1}{1-\zeta}}}{\sigma^{\frac{1}{1-\zeta}}} \left(w_o^R\right)^{\frac{\zeta}{\zeta-1}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}}, \\
&= \left(\zeta^{\frac{\zeta}{1-\zeta}} - \zeta^{\frac{1}{1-\zeta}} \right) \left(w_o^R\right)^{\frac{\zeta}{\zeta-1}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\frac{1}{\sigma} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}}, \\
&= \kappa_\zeta^{\frac{1}{1-\zeta}} \left(w_o^{R^{-\zeta}} \psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \frac{1}{\sigma} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}}, \\
&= \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R^\zeta} w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}.
\end{aligned}
$$

where $\kappa_\zeta \equiv (1-\zeta)^{1-\zeta}\,\zeta^\zeta$ is a normalization constant. The normalization constant is

$$
\begin{aligned}
\zeta^{\frac{\zeta}{1-\zeta}} - \zeta^{\frac{1}{1-\zeta}} &= \zeta^{\frac{\zeta}{1-\zeta}} - \zeta^{\frac{1}{1-\zeta}}, \\
&= \zeta^{\frac{1}{1-\zeta}} \left[\zeta^{\frac{\zeta}{1-\zeta}-\frac{1}{1-\zeta}} - 1\right], \\
&= \zeta^{\frac{1}{1-\zeta}} \left[\zeta^{\frac{\zeta-1}{1-\zeta}} - 1\right], \\
&= \zeta^{\frac{1}{1-\zeta}} \left[\zeta^{-\frac{1-\zeta}{1-\zeta}} - 1\right], \\
&= \zeta^{\frac{1}{1-\zeta}} \left[\frac{1}{\zeta} - 1\right], \\
&= \zeta^{\frac{1}{1-\zeta}} \left[\frac{1-\zeta}{\zeta}\right], \\
&= \zeta^{\frac{1}{1-\zeta}-1} (1-\zeta), \\
&= \zeta^{\frac{1-1+\zeta}{1-\zeta}} (1-\zeta), \\
&= (1-\zeta)\,\zeta^{\frac{\zeta}{1-\zeta}}, \\
&= \left[(1-\zeta)\,\zeta^{\frac{\zeta}{1-\zeta}}\right]^{\frac{1-\zeta}{1-\zeta}}, \\
&= \left[(1-\zeta)^{1-\zeta}\,\zeta^{\zeta}\right]^{\frac{1}{1-\zeta}}, \\
&= \kappa_\zeta^{\frac{1}{1-\zeta}}.
\end{aligned}
$$

R&D with subsidies. Firm's demand for inventors is

$$
R_o^\omega = \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{(1-s_o)\, w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{1}{1-\zeta}}.
$$

The number of ideas is

$$
n_o^\omega = \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{(1-s_o)\, w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}.
$$

Intermediate's quality is

$$
A_o^\omega = \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{\zeta}}}{(1-s_o)\, w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{\zeta}{1-\zeta}}.
$$

Firms' profits are

$$\overline{\pi}_o^{\omega} = \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{(1-s_o)^{\zeta} w_o^{R\zeta} w_o^{L\sigma-1}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}.$$

Preferences. In each location d, agents are of two types: inventors ($n = R$), or workers ($n = L$). Each agent has preferences over the local final good and location amenities:

$$\begin{aligned} &\max_{\{Q_d^n\}} \quad U_d^n = \mathcal{B}_d^n Q_d^n \quad s.t. \\ &w_d^n (1 + \overline{\pi}) = P_d Q_d^n, \end{aligned}$$

where $\mathcal{B}_d^n$ are type-specific location amenities, Q_d^n is the quantity demanded by agent of type n, w_d^n is the agent's wage, and $\overline{\pi}$ are redistributed profits. Since the agent's preferences are linear, utility is maximized at

$$\begin{aligned} U_d^n &= \mathcal{B}_d^n Q_d^n, \\ &= \mathcal{B}_d^n \left(\frac{w_d^n (1+\overline{\pi})}{P_d} \right), \\ &= \frac{\mathcal{B}_d^n w_d^n (1+\overline{\pi})}{P_d}. \end{aligned}$$

Location choice. An agent i of type $n = \{L, R\}$ living in o moves to d by maximizing its *ex ante* indirect utility:

$$U_{od}^{i,n} = \max_{d \in \mathcal{S}} \left\{ \frac{U_d^n}{\mu_{od}^n} \times \epsilon^i \right\},$$

where $\mu_{od}^n \geq 1$ are type-specific *iceberg* migration costs, $G(\epsilon) = \exp(-\epsilon^{-\kappa})$ are location preference shocks, and κ is the spatial labor supply elasticity. Following the properties of the Frechet distribution, the share of agents of type n moving from o to d is

$$\eta_{od}^n = \frac{\left(\frac{U_d^n}{\mu_{od}^n}\right)^{\kappa}}{\sum_\delta \left(\frac{U_\delta^n}{\mu_{o\delta}^L}\right)^{\kappa}},$$

such that $\sum_d \eta_{od}^n = 1, \forall o \in \mathcal{S}$.

Aggregate productivity. I define aggregate productivity as the average quality of intermediates in a location. Since firms are symmetric and the mass of firms in each location is fixed, then from Equations (1.14)-(1.15), location's productivity is

$$
\begin{aligned}
A_o &= \int_{\omega \in \Omega_o} A_o^\omega d\omega, \\
&= \int_{\omega \in \Omega_o} \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{\zeta}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{\zeta}{1-\zeta}} d\omega, \\
&= \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{\zeta}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{\zeta}{1-\zeta}} \int_{\omega \in \Omega_o} d\omega, \\
&= \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{\zeta}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \right)^{\zeta} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right) \left(w_o^{L^{\sigma-1}} w_o^R \right)^{-\zeta} \left(\sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\zeta} \right)^{\frac{1}{1-\zeta}}.
\end{aligned}
$$

With R&D subsidies, a the productivity of a location is

$$
A_o = \left(\left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \right)^{\zeta} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right) \left((1 - s_o) \, w_o^{L^{\sigma-1}} w_o^R \right)^{-\zeta} \left(\sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\zeta} \right)^{\frac{1}{1-\zeta}}.
$$

Then, aggregate productivity is

$$
A = \frac{1}{S} \sum_o A_o,
$$

where $S \equiv |\mathcal{S}|$ is the number of locations in the economy.

Price indices. Given Equation (1.9), the price index of firms in o selling to d is

$$
\begin{aligned}
P_{od}^{1-\sigma} &= \int_{\omega \in \Omega_{od}} A_o^\omega P_{od}^{\omega^{1-\sigma}} d\omega, \\
&= \int_{\omega \in \Omega_{od}} A_o^\omega \left(\overline{m} \tau_{od} w_o^L \right)^{1-\sigma} d\omega, \\
&= \int_{\omega \in \Omega_{od}} A_o^\omega d\omega \left(\overline{m} \tau_{od} w_o^L \right)^{1-\sigma}, \\
&= \int_{\omega \in \Omega_o} A_o^\omega d\omega \left(\overline{m} \tau_{od} w_o^L \right)^{1-\sigma}, \\
&= A_o \left(\overline{m} \tau_{od} w_o^L \right)^{1-\sigma}.
\end{aligned}
$$

Then, the price index in d is

$$
\begin{aligned}
P_d^{1-\sigma} &= \sum_o P_{od}^{1-\sigma}, \\
&= \sum_o A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}.
\end{aligned}
$$

Trade shares. From (1.5), trade flows from o to d are

$$
\begin{aligned}
Q_{od}^{\omega} &= A_o^{\omega} P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d, \\
P_{od}^{\omega} Q_{od}^{\omega} &= A_o^{\omega} P_{od}^{\omega^{1-\sigma}} P_d^{\sigma-1} X_d, \\
X_{od}^{\omega} &= A_o^{\omega} P_{od}^{\omega^{1-\sigma}} P_d^{\sigma-1} X_d, \\
\int_{\omega\in\Omega_{od}} X_{od}^{\omega} d\omega &= \int_{\omega\in\Omega_{od}} A_o^{\omega} P_{od}^{\omega^{1-\sigma}} P_d^{\sigma-1} X_d d\omega, \\
X_{od} &= \left(\int_{\omega\in\Omega_{od}} A_o^{\omega} P_{od}^{\omega^{1-\sigma}} d\omega\right) P_d^{\sigma-1} X_d, \\
&= P_{od}^{1-\sigma} P_d^{\sigma-1} X_d,
\end{aligned}
$$

where $P_{od}^{1-\sigma} = \int_{\omega\in\Omega_{od}} A_o^{\omega} P_{od}^{\omega^{1-\sigma}} d\omega$. Then, the share of intermediate inputs from o in location d's expenditure χ_{od} is

$$
\begin{aligned}
\chi_{od} &\equiv \frac{X_{od}}{X_d}, \\
&= \frac{P_{od}^{1-\sigma} P_d^{\sigma-1} X_d}{X_d}, \\
&= P_{od}^{1-\sigma} P_d^{\sigma-1}, \\
&= \frac{P_{od}^{1-\sigma}}{P_d^{1-\sigma}}.
\end{aligned}
$$

Considering Equations (1.22), then trade shares are

$$
\begin{aligned}
\chi_{od} &= \frac{P_{od}^{1-\sigma}}{P_d^{1-\sigma}}, \\
&= \frac{A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}}{\sum_o A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}}, \\
&= \frac{A_o \left(\tau_{od} w_o^L\right)^{1-\sigma}}{\sum_o A_o \left(\tau_{od} w_o^L\right)^{1-\sigma}}.
\end{aligned}
$$

Profits per-capita. Location's profits are

$$
\begin{aligned}
\overline{\pi}_o &= \int_{\omega\in\Omega_o} \overline{\pi}_o^{\omega} d\omega, \\
&= \int_{\omega\in\Omega_o} \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R\zeta} w_o^{L\sigma-1}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}} d\omega, \\
&= \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R\zeta} w_o^{L\sigma-1}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}} \int_{\omega\in\Omega_o} d\omega, \\
&= \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma} \psi}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R\zeta} w_o^{L\sigma-1}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}.
\end{aligned}
$$

With R&D subsidies, these are

$$
\overline{\pi}_o = \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma} \psi}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}}}{(1-s_o)^{\zeta}\, w_o^{R\zeta} w_o^{L\sigma-1}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}.
$$

Then, profits per-capita are

$$
\overline{\pi} = \frac{1}{N} \sum_o \pi_o,
$$

where $N = L + R$ is total population.

Workers market. From (1.5), (1.7), and (1.9), the demand for workers is

$$
\begin{aligned}
L_o &= \sum_d \int_{\omega \in \Omega_{od}} L_{od}^{\omega} d\omega, \\
&= \sum_d \int_{\omega \in \Omega_{od}} Q_{od}^{\omega} d\omega, \\
&= \sum_d \int_{\omega \in \Omega_{od}} \left(A_o^{\omega} P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d \right) d\omega, \\
&= \sum_d \int_{\omega \in \Omega_{od}} \left(A_o^{\omega} \left(\overline{m} \tau_{od} w_o^L \right)^{-\sigma} P_d^{\sigma-1} X_d \right) d\omega, \\
&= \sum_d \left(\overline{m} \tau_{od} w_o^L \right)^{-\sigma} P_d^{\sigma-1} X_d \int_{\omega \in \Omega_{od}} A_o^{\omega} d\omega, \\
&= \sum_d \left(\overline{m} \tau_{od} w_o^L \right)^{-\sigma} P_d^{\sigma-1} X_d \int_{\omega \in \Omega_{o}} A_o^{\omega} d\omega, \\
&= \sum_d \left(\overline{m} \tau_{od} w_o^L \right)^{-\sigma} P_d^{\sigma-1} X_d A_o, \\
&= A_o \left(\overline{m} w_o^L \right)^{-\sigma} \sum_d \tau_{od}^{-\sigma} P_d^{\sigma-1} X_d, \\
w_o^{L^{\sigma}} &= \frac{A_o}{L_o} \overline{m}^{-\sigma} \sum_d \tau_{od}^{-\sigma} P_d^{\sigma-1} X_d, \\
w_o^L &= \overline{m}^{-1} \left(\frac{A_o}{L_o} \sum_d \tau_{od}^{-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{\sigma}}.
\end{aligned}
$$

Inventors market. From firm's demand for inventors (1.17), location's demand for inventors is

$$
\begin{aligned}
R_o^\omega &= \left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma}\frac{\psi\mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R w_o^{L^{\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d\right)^{\frac{1}{1-\zeta}}, \\
\int_{\omega\in\Omega_o} R_o^\omega d\omega &= \int_{\omega\in\Omega_o}\left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma}\frac{\psi\mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R w_o^{L^{\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d\right)^{\frac{1}{1-\zeta}} d\omega, \\
R_o &= \left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma}\frac{\psi\mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R w_o^{L^{\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d\right)^{\frac{1}{1-\zeta}} \int_{\omega\in\Omega_o} d\omega, \\
&= R_o^{\frac{\widetilde{\gamma}}{1-\zeta}}\left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma}\frac{\psi\mathcal{A}_o}{w_o^R w_o^{L^{\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d\right)^{\frac{1}{1-\zeta}}, \\
R_o^{1-\frac{\widetilde{\gamma}}{1-\zeta}} &= \left(\frac{\zeta\psi\overline{m}^{1-\sigma}}{\sigma}\frac{\mathcal{A}_o}{w_o^R w_o^{L^{\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d\right)^{\frac{1}{1-\zeta}}, \\
R_o^{\frac{(1-\zeta)-\widetilde{\gamma}}{1-\zeta}} &= \left(\frac{\zeta\psi\overline{m}^{1-\sigma}}{\sigma}\frac{\mathcal{A}_o}{w_o^R w_o^{L^{\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d\right)^{\frac{1}{1-\zeta}}, \\
R_o^{(1-\zeta)-\widetilde{\gamma}} &= \frac{\zeta\psi\overline{m}^{1-\sigma}}{\sigma}\frac{\mathcal{A}_o}{w_o^R w_o^{L^{\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d, \\
w_o^R &= \frac{\zeta\psi\overline{m}^{1-\sigma}}{\sigma}\frac{\mathcal{A}_o R_o^{\widetilde{\gamma}-(1-\zeta)}}{w_o^{L^{\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d.
\end{aligned}
$$

With R&D subsidies, this is

$$
w_o^R = \frac{\zeta\psi\overline{m}^{1-\sigma}}{\sigma}\frac{\mathcal{A}_o R_o^{\widetilde{\gamma}-(1-\zeta)}}{(1-s_o)\,w_o^{L^{\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d.
$$

A.6 Taking the Model to the Data

Spatial knowledge spillovers in R&D $\{\widetilde{\gamma}\}$. From Equation (1.13), inventor productivity can be written as

$$
\begin{aligned}
\mathbb{E}\left\{Z_o^{i\omega}\right\} &= \psi\lambda_o^{\frac{1}{\alpha}}, \\
&= \psi\mathcal{A}_o R_o^{\widetilde{\gamma}}, \\
Z_o^{i\omega} &= \psi\mathcal{A}_o R_o^{\widetilde{\gamma}} \exp\left(\epsilon_o^{i\omega}\right), \\
\log\left(Z_o^{i\omega}\right) &= \log\left(\psi\mathcal{A}_o R_o^{\widetilde{\gamma}} \exp\left(\epsilon_o^{i\omega}\right)\right), \\
&= \log\left(\psi\right) + \log\left(\mathcal{A}_o\right) + \log\left(R_o^{\widetilde{\gamma}}\right) + \log\left(\exp\left(\epsilon_o^{i\omega}\right)\right), \\
&= \log\left(\psi\right) + \log\left(\mathcal{A}_o\right) + \widetilde{\gamma}\log\left(R_o\right) + \epsilon_o^{i\omega}, \\
&= \iota + \iota_o + \widetilde{\gamma}\log\left(R_o\right) + \epsilon_o^{i\omega},
\end{aligned}
$$

where $\iota \equiv \log\left(\psi\right)$ and $\iota \equiv \log\left(\mathcal{A}_o\right)$. After considering the additional time dimension t and technological areas a, and first differences, Equation (1.34) is the model counterpart of Equation (1.2) which is used to estimate spatial knowledge spillovers in R&D $\beta = 0.409$.

Migration costs $\{\mu_{od}^n\}$. Migration costs are parametrized by data on geographic distance between every location pair $o, d \in \mathcal{S}, \mathcal{S}$, intercepts $\left\{\rho_0^L, \rho_0^R\right\}$, and elasticities of migration costs to distance $\left\{\rho_1^L, \rho_1^R\right\}$. The intercepts are calibrated by targeting the overall migration rate for workers and inventors. To estimate the migration cost elasticities, consider the share of inventors in Equation (1.20), such that

$$
\begin{aligned}
\eta_{od}^{n} &= \frac{\left(\frac{U_d^n}{\mu_{od}^n}\right)^{\kappa}}{\sum_{\delta}\left(\frac{U_\delta^n}{\mu_{o\delta}^R}\right)^{\kappa}}, \\
\log\left(\eta_{od}^{n}\right) &= \log\left(\frac{\left(\frac{U_d^n}{\mu_{od}^n}\right)^{\kappa}}{\sum_{\delta}\left(\frac{U_{\delta,t}^n}{\mu_{o\delta}^n}\right)^{\kappa}}\right), \\
&= \log\left(\left(\frac{U_d^n}{\mu_{od}^n}\right)^{\kappa}\right) - \log\left(\sum_{\delta}\left(\frac{U_\delta^n}{\mu_{o\delta}^n}\right)^{\kappa}\right), \\
&= \kappa\log\left(\frac{U_d^n}{\mu_{od}^n}\right) - \log\left(\sum_{\delta}\left(\frac{U_\delta^n}{\mu_{o\delta}^n}\right)^{\kappa}\right), \\
&= -\kappa\log\left(\mu_{od}^n\right) \underbrace{- \log\left(\sum_{\delta}\left(\frac{U_\delta^R}{\mu_{o\delta}^n}\right)^{\kappa}\right)}_{=\iota_o} + \underbrace{\kappa\log\left(U_d^n\right)}_{=\iota_d}, \\
&= \iota_o + \iota_d - \kappa\log\left(\mu_{od}^n\right), \\
&= \iota_o + \iota_d - \kappa\log\left(\rho_0^n dist_{od}^{\rho_1^n}\exp\left(-\frac{\epsilon_{od}^n}{\kappa}\right)\right), \\
&= \iota_o + \iota_d \underbrace{-\kappa\log\left(\rho_0^n\right)}_{\iota} - \kappa\log\left(dist_{od}^{\rho_1^n}\right) - \kappa\log\left(\exp\left(-\frac{\epsilon_{od}^n}{\kappa}\right)\right), \\
&= \iota + \iota_o + \iota_d - \kappa\rho_1^n\log\left(dist_{od}\right) + \epsilon_{od}^n.
\end{aligned}
$$

This migration gravity equation states that, conditional on data on migration shares $\{\eta_{od}^n\}$, geographic distances $\{dist_{od}\}$, the migration elasticity $\{\kappa\}$, and the inclusion of origin and destination fixed effects $\{\iota_o, \iota_d\}$, then migration cost elasticities $\{\rho_1^n\}$ are identified in the data.

Fundamental location productivity $\{\mathcal{A}_o\}$. Unobserved fundamental location productivities are recovered through model inversion. Given parameters $\{\sigma, \widetilde{\gamma}\}$, trade costs $\{\tau_{od}\}_{\forall o,d\in\mathcal{S},\mathcal{S}}$, and data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o\in\mathcal{S}}$, there is a unique set of values for fundamental location productivities $\{\mathcal{A}_o\}_{\forall o\in\mathcal{S}}$ that is consistent with the data. Given equilibrium in goods market (1.32), trade shares (1.23), and aggregate productivity (1.21), I construct the following system of excess demand functions:

$$\mathbb{D}_o(\mathcal{A}) \equiv w_o^L L_o + w_o^R R_o - \sum_d \chi_{od}\left(w_d^L L_d + w_d^R R_d\right),$$
$$= w_o^L L_o + w_o^R R_o - \sum_d \frac{A_o\left(\tau_{od} w_o^L\right)^{1-\sigma}}{\sum_o A_o\left(\tau_{od} w_o^L\right)^{1-\sigma}}\left(w_d^L L_d + w_d^R R_d\right),$$

where A_o is a function of location fundamentals $\mathcal{A}_o$. It can be shown that this excess demand functions are (i) continuous, (ii) homogeneous of degree zero, (iii) $\sum_o \mathbb{D}_o(\mathcal{A}) = 0$, and (iv) $\frac{\partial \mathbb{D}_o(\mathcal{A})}{\partial \mathcal{A}_l} > 0, \forall o, l \in \mathcal{S}, \mathcal{S}, l \neq o$ and $\frac{\partial \mathbb{D}_o(\mathcal{A})}{\partial \mathcal{A}_o} < 0, \forall o \in \mathcal{S}$. Given this properties, up to a normalization, there exists a unique vector $\mathcal{A}^*$ such that $\mathbb{D}_o(\mathcal{A}^*) = 0, \forall o \in \mathcal{S}$. I use data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o \in \mathcal{S}}$ for year 2014.

Fundamental location amenities $\{\mathcal{B}_o^n\}$. Unobserved fundamental location amenities are recovered through model inversion. Given parameters $\{\sigma, \kappa, \widetilde{\gamma}\}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, migration costs $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, fundamental location productivities $\{\mathcal{A}_o\}_{\forall o \in \mathcal{S}}$, and data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o \in \mathcal{S}}$, there is a unique set of values for fundamental location amenities $\{\mathcal{B}_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$ that is consistent with the data. Given labor supply functions (1.27) and (1.29), migration shares (1.20), and indirect utility functions (1.18), I construct the following system of excess demand functions:

$$\mathbb{D}_d^R\left(\mathcal{B}^R\right) = R_d - \sum_o \eta_{od}^R \overline{R}_o,$$
$$= R_d - \sum_o \left(\frac{\left(\frac{U_d^R}{\mu_{od}^R}\right)^\kappa}{\sum_\delta \left(\frac{U_\delta^R}{\mu_{o\delta}^R}\right)^\kappa}\right)\overline{R}_o,$$
$$= R_d - \sum_o \left(\frac{\left(\frac{\mathcal{B}_d^R w_d^R}{\mu_{od}^R P_d}\right)^\kappa}{\sum_\delta \left(\frac{\mathcal{B}_\delta^R w_\delta^R}{\mu_{o\delta}^R P_\delta}\right)^\kappa}\right)\overline{R}_o.$$

The same procedure can be applied for workers:

$$\mathbb{D}_d^L\left(\mathcal{B}^L\right) = L_d - \sum_o \left(\frac{\left(\frac{\mathcal{B}_d^L w_d^L}{\mu_{od}^L P_d}\right)^\kappa}{\sum_\delta \left(\frac{\mathcal{B}_\delta^L w_\delta^L}{\mu_{o\delta}^L P_\delta}\right)^\kappa}\right)\overline{L}_o.$$

Prices $\{P_d\}_{\forall d \in \mathcal{S}}$ are constructed given Equations (1.22) and (1.21). It can be shown that these excess demand functions are (i) continuous, (ii) homogeneous of degree zero, (iii) $\sum_o \mathbb{D}_d^n(\mathcal{B}^n) = 0$, and (iv) $\frac{\partial \mathbb{D}_d^n(\mathcal{B}^n)}{\partial \mathcal{B}_l^n} > 0, \forall d, l \in \mathcal{S}, \mathcal{S}, l \neq o$ and $\frac{\partial \mathbb{D}_d^n(\mathcal{B}^n)}{\partial \mathcal{B}_d^n} < 0, \forall d \in \mathcal{S}$. Given this properties, up to a normalization, there exists a unique vector $\mathcal{B}^{n^*}$ such that $\mathbb{D}_d^n\left(\mathcal{B}^{n^*}\right) = 0, \forall d \in \mathcal{S}, n = \{L, R\}$. I use data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o \in \mathcal{S}}$ for year 2014.

A.7 Solution algorithms

In this section I describe the algorithm that solves the model. The supra-script (i) denotes a variable as an "input", and the supra-script (o) denotes a variable as an "output".

A.7.1 Equilibrium

Given the exogenous distribution of workers and inventors across locations $\left\{\overline{L}_o, \overline{R}_o\right\}_{\forall o \in \mathcal{S}}$, location fundamentals $\left\{\mathcal{A}_o, \mathcal{B}_o^L, \mathcal{B}_o^R\right\}_{\forall o \in \mathcal{S}}$, migration costs $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, and parameters, the model is solved following these steps:

1. Guess $\left\{w_o^{L^{(i)}}, w_o^{R^{(i)}}, A_o^{(i)}\right\}_{\forall o \in \mathcal{S}}$ and $\overline{\pi}^{(i)}$:
   (a) Bilateral price indices $\{P_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$:
   $$P_{od}^{1-\sigma} = A_o^{(i)}\left(\overline{m}\tau_{od} w_o^{L^{(i)}}\right)^{1-\sigma}$$
   (b) Price indices $\{P_d\}_{\forall d \in \mathcal{S}}$:
   $$P_d = \left(\sum_o P_{od}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$$
   (c) Migration shares $\{\eta_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$:
   $$\eta_{od}^n = \frac{\left(\frac{\mathcal{B}_d^n w_d^{n^{(i)}}}{\mu_{od}^n P_d}\right)^\kappa}{\sum_\delta \left(\frac{\mathcal{B}_\delta^n w_\delta^{n^{(i)}}}{\mu_{o\delta}^n P_\delta}\right)^\kappa}$$
   (d) Number of workers and inventors $\{L_d, R_d\}_{\forall d \in \mathcal{S}}$:
   $$L_d = \sum_o \eta_{od}^L \overline{L}_o,$$
   $$R_d = \sum_o \eta_{od}^R \overline{R}_o$$
   (e) Income $\{Y_o\}_{\forall o \in \mathcal{S}}$:
   $$Y_o = \left(1 + \overline{\pi}^{(i)}\right)\left(w_o^{L^{(i)}} L_o + w_o^{R^{(i)}} R_o\right)$$

(f) Expenditure equals income:

$$X_o = Y_o$$

(g) Location profits $\{\overline{\pi}_o\}_{\forall o \in \mathcal{S}}$:

$$\overline{\pi}_o = \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}\psi}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R^{(i)}\zeta} w_o^{L^{(i)}\sigma-1}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}$$

(h) Profits per-capita $\overline{\pi}^{(o)}$:

$$\overline{\pi}^{(o)} = \frac{1}{N} \sum_o \overline{\pi}_o,$$

where $N = L + R$ is total population

(i) New worker wages $\left\{ w_o^{L^{(o)}} \right\}_{\forall o \in \mathcal{S}}$:

$$w_o^{L^{(o)}} = \overline{m}^{-1} \left(\frac{A_o^{(i)}}{L_o} \sum_d \tau_{od}^{-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{\sigma}}$$

(j) New inventor wages $\left\{ w_o^{R^{(o)}} \right\}_{\forall o \in \mathcal{S}}$:

$$w_o^{R^{(o)}} = \frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}-(1-\zeta)}}{w_o^{L^{(i)}\sigma-1}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d$$

(k) Normalize wages such that $w_1^{L^{(o)}} = 1$

(l) New location productivity $\left\{ A_o^{(o)} \right\}_{\forall o \in \mathcal{S}}$:

$$A_o^{(o)} = \left(\left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \right)^\zeta \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right) \left(w_o^{L^{(i)}\sigma-1} w_o^{R^{(i)}} \right)^{-\zeta} \left(\sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^\zeta \right)^{\frac{1}{1-\zeta}}$$

(m) Update $w_o^{R^{(i)}} = w_o^{R^{(o)}}$, $w_o^{L^{(i)}} = w_o^{L^{(o)}}$, $A_o^{(i)} = A_o^{(o)}$ and $\overline{\pi}^{(i)} = \overline{\pi}^{(o)}$

(n) Iterate until convergence is achieved

A.7.2 Equilibrium with R&D subsidies

Given the exogenous distribution of workers and inventors across locations $\left\{ \overline{L}_o, \overline{R}_o \right\}_{\forall o \in \mathcal{S}}$, location fundamentals $\left\{ \mathcal{A}_o, \mathcal{B}_o^L, \mathcal{B}_o^R \right\}_{\forall o \in \mathcal{S}}$, migration costs $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, R&D subsidies $\{s_o\}_{\forall o \in \mathcal{S}}$, and parameters, the model is solved following these steps:

1. Guess $\left\{w_o^{L(i)}, w_o^{R(i)}, A_o^{(i)}\right\}_{\forall o \in \mathcal{S}}$ and $\left\{\overline{\pi}^{(i)}, \tau^{(i)}\right\}$:

   (a) Bilateral price indices $\{P_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$:

$$P_{od}^{1-\sigma} = A_o^{(i)} \left(\overline{m}\tau_{od} w_o^{L(i)}\right)^{1-\sigma}$$

   (b) Price indices $\{P_d\}_{\forall d \in \mathcal{S}}$:

$$P_d = \left(\sum_o P_{od}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$$

   (c) Migration shares $\{\eta_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$:

$$\eta_{od}^n = \frac{\left(\frac{\mathcal{B}_d^n w_d^{n(i)}}{\mu_{od}^n P_d}\right)^\kappa}{\sum_\delta \left(\frac{\mathcal{B}_\delta^n w_\delta^{n(i)}}{\mu_{o\delta}^n P_\delta}\right)^\kappa}$$

   (d) Number of workers and inventors $\{L_d, R_d\}_{\forall d \in \mathcal{S}}$:

$$L_d = \sum_o \eta_{od}^L \overline{L}_o,$$
$$R_d = \sum_o \eta_{od}^R \overline{R}_o$$

   (e) Income $\{Y_o\}_{\forall o \in \mathcal{S}}$:

$$Y_o = \left(1 + \overline{\pi}^{(i)} + \tau^{(i)}\right)\left(w_o^{L(i)} L_o + w_o^{R(i)} R_o\right)$$

   (f) Expenditure equals income:

$$X_o = Y_o$$

   (g) Location profits $\{\overline{\pi}_o\}_{\forall o \in \mathcal{S}}$:

$$\overline{\pi}_o = \left(\frac{\kappa\zeta\overline{m}^{1-\sigma}\psi}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}}}{(1-s_o)^\zeta\, w_o^{R(i)\zeta}\, w_o^{L(i)\sigma-1}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d\right)^{\frac{1}{1-\zeta}}$$

   (h) Profits per-capita $\overline{\pi}^{(o)}$:

$$\overline{\pi}^{(o)} = \frac{1}{N}\sum_o \overline{\pi}_o,$$

   where $N = L + R$ is total population

(i) New value of tax rate $\tau^{(o)}$:

$$\tau^{(o)} = \frac{\sum_o s_o \left(w_o^{R^{(i)}} R_o \right)}{\sum_o \left(w_o^{L^{(i)}} L_o + w_o^{R^{(i)}} R_o \right)}$$

(j) New worker wages $\left\{ w_o^{L^{(o)}} \right\}_{\forall o \in \mathcal{S}}$:

$$w_o^{L^{(o)}} = \overline{m}^{-1} \left(\frac{A_o^{(i)}}{L_o} \sum_d \tau_{od}^{-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{\sigma}}$$

(k) New inventor wages $\left\{ w_o^{R^{(o)}} \right\}_{\forall o \in \mathcal{S}}$:

$$w_o^{R^{(o)}} = \frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}-(1-\zeta)}}{(1-s_o)\, w_o^{L^{(i)}\sigma-1}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d$$

(l) Normalize wages such that $w_1^{L^{(o)}} = 1$

(m) New location productivity $\left\{ A_o^{(o)} \right\}_{\forall o \in \mathcal{S}}$:

$$A_o^{(o)} = \left(\left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \right)^{\zeta} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right) \left((1-s_o)\, w_o^{L^{(i)}\sigma-1} w_o^{R^{(i)}} \right)^{-\zeta} \left(\sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\zeta} \right)^{\frac{1}{1-\zeta}}$$

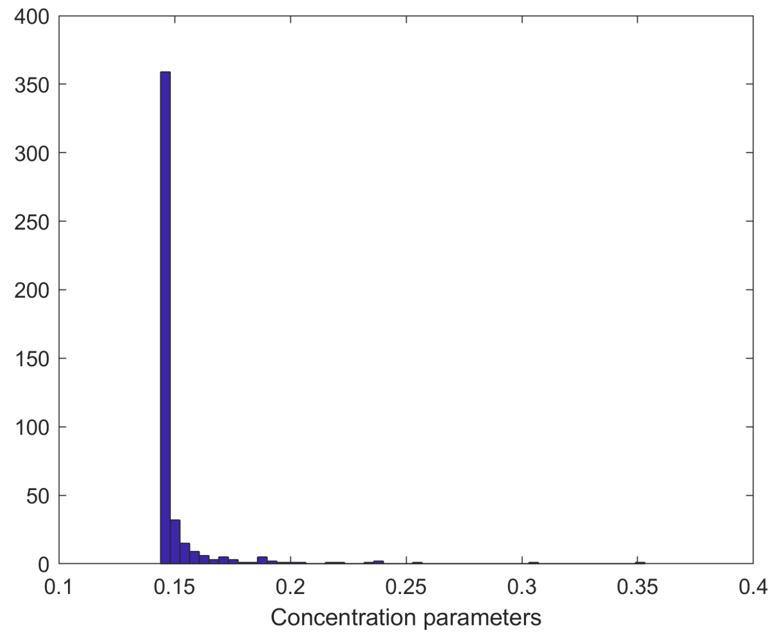
(n) Update $w_o^{R^{(i)}} = w_o^{R^{(o)}}$, $w_o^{L^{(i)}} = w_o^{L^{(o)}}$, $A_o^{(i)} = A_o^{(o)}$, $\overline{\pi}^{(i)} = \overline{\pi}^{(o)}$, and $\tau^{(o)} = \tau^{(i)}$

(o) Iterate until convergence is achieved

Appendices to Chapter II

B.1 Additional figures and tables

Figure B1: Histogram of estimated concentration parameters for Dirichlet distribution



Notes: Estimated concentration parameters for a Dirichlet distribution according to the maximum likelihood estimation from Equation (2.16).

Table B1: Estimation for matching cost

	(1)	(2)
	1st Stage	2nd Stage
Dep. Variable	Sales (Hyperbolic Inverse Sine)	Trade Indicator
BC	0.013***	0.131***
	(0.001)	(0.008)
$\text{ihs}\,\widehat{[n\,(z,z')]}$		8.340***
		(0.024)
Obs.	5,606,627	5,606,627
Adj. R2	0.595	-
Pseudo R2	-	0.453
FE	Seller, buyer	-

Notes: Column 1 shows the results of estimating Equation (2.17). Column 2 shows the results of estimating Equation (2.18). We winsorizeln$\widehat{n\,(z,z')}$ at 1 percent and 99 percent. Sample only contains in-state firms. ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Standard errors clustered at the seller and buyer level in Column 1. Standard errors in parentheses. The higher the Bhattacharyya coefficient, the culturally closer two firms are.

B.2 Targeted and untargeted moments

B.2.1 Normalized number of buyers and sellers

Data. In our dataset, for each firm i, we calculate the number firms it sold to and the number of firms it bought from. Then, to normalize this measure, we divide this number by the total number of firms in our sample. Thus, for a specific firm i, we can understand this measure as the share of firms this specific firm i is connected to, both as a buyer and a seller.

Model. For this part we start with the link function matrix, where each element $l\left(z, z^{'}\right)$ represents the pairwise probability that seller z will match with buyer $z^{'}$. For each seller z, we take the average $l\left(z, z^{'}\right)$ across all the possible buyers. This represents the proportion of firms that seller z will match to with respect to the total number of firms. We multiply this number by the total number of firms $\mathcal{N}$to obtain the number of buyers for each seller z. We follow a similar procedure to calculate the number of sellers each buyer $z^{'}$ has.

B.2.2 Normalized intermediate sales and purchases

Data. In our dataset, for each firm i, we calculate the total sales to other firms and the total purchases from other firms. In the case of the sellers, we normalize this measure by dividing the total sales of firm i by the total number of buyers this firm has. We follow a similar procedure with the buyers to calculate the normalized intermediate purchases.

Model. We use the intermediate sales matrix, where each element $n\left(z, z^{'}\right)$ represents the total sales of intermediate goods from seller z to buyer $z^{'}$. We sum all the sales for each seller z and divide this number by the number of buyers it has. Thus, we obtain the normalized intermediate sales for a given seller. For the normalized intermediate purchases we follow a similar procedure with the buyers.

Table B2: Targeted and untargeted moments

Targeted Moments		
	Data	Model
$mean\left[\ln\left(\mathcal{N}_b\left(\nu\right)/\mathcal{N}\right)\right]$	-9.24	-9.48
$var\left[\ln\left(\mathcal{N}_b\left(\nu\right)/\mathcal{N}\right)\right]$	0.98	0.89
$var\left[\ln\left(\widetilde{N}\left(\nu\right)/\mathcal{N}_b\left(\nu\right)\right)\right]$	2.82	2.82
$mean\left[\ln\left(\mathcal{N}_s\left(\omega\right)/\mathcal{N}\right)\right]$	-9.39	-9.14
Untargeted Moments		
	Data	Model
$var\left[\ln\left(\mathcal{N}_s\left(\omega\right)/\mathcal{N}\right)\right]$	0.60	0.16
$var\left[\ln\left(N\left(\omega\right)/\mathcal{N}_s\left(\omega\right)\right)\right]$	2.73	0.56

Notes: The targeted moments are the mean of the log-normalized number of buyers $mean\left[\ln\left(\mathcal{N}_b\left(\nu\right)/\mathcal{N}\right)\right]$, the variance of the log-normalized number of buyers $var\left[\ln\left(\mathcal{N}_b\left(\nu\right)/\mathcal{N}\right)\right]$ and the variance of the log-normalized intermediate sales $var\left[\ln\left(\widetilde{N}\left(\nu\right)/\mathcal{N}_b\left(\nu\right)\right)\right]$, where $\widetilde{N}\left(\nu\right)$ are the total intermediate sales of seller ν. The untargeted moments are the mean of the log-normalized number of sellers $mean\left[\ln\left(\mathcal{N}_s\left(\omega\right)/\mathcal{N}\right)\right]$, the variance of the log-normalized number of sellers $var\left[\ln\left(\mathcal{N}_s\left(\omega\right)/\mathcal{N}\right)\right]$ and the variance of the log-normalized intermediate purchases $var\left[\ln\left(N\left(\omega\right)/\mathcal{N}_s\left(\omega\right)\right)\right]$.

B.3 Additional specifications

B.3.1 Kullback-Leibler divergence

In this section we present an alternative measure of cultural proximity to that of the Bhattacharyya coefficient. Define the standard discrete distribution-based Kullback and Leibler (1951) divergence as

$$KL\left(\nu\|\omega\right) = \sum_{x=1}^{X} \rho_{\nu}\left(x\right) \log\left(\frac{\rho_{\nu}\left(x\right)}{\rho_{\omega}\left(x\right)}\right).$$

We have that $KL\left(\nu\|\omega\right) \geq 0$, where $KL\left(\nu\|\omega\right) = 0$ when sellers and buyers have exactly equal probability distributions, while it will be higher the more different the two probability distributions are.[47] Intuitively, we can see this measure as the expected difference between two probability distributions. However, this proximity measure is not symmetric; that is, $KL\left(\nu\|\omega\right) \neq KL\left(\omega\|\nu\right)$. Consider our previous example where we record a transaction between a seller ν and a buyer with distribution ω, from which we calculate $KL\left(\nu\|\omega\right)$. If, in a second transaction, the roles of the firms revert, then the Kullback-Leibler divergence would be $KL\left(\omega\|\nu\right)$, implying the cultural proximity between the two firms has changed, when it should not change. To convert this measure into a symmetric one, we define

$$KL_{sym}\left(\nu\|\omega\right) = KL\left(\nu\|\omega\right) + KL\left(\omega\|\nu\right) = KL_{sym}\left(\omega\|\nu\right).$$

Notice this similarity measure needs $\rho_{\nu}\left(x\right) > 0$ and $\rho_{\omega}\left(x\right) > 0$ for all x. However, it is possible that the probability of a firm belonging to a certain cultural group is zero. In those cases we replace that probability of zero for a probability $\varepsilon \to 0^{+}$ such that KL_{sym} is well-defined. Tables B3 and B4 show the regression results for the intensive margin, unit prices and extensive margin, respectively. In this case, the higher the Kullback-Leibler divergence, the more culturally different the buyer from the seller. The results confirm the findings from the main text.

[47] This interpretation diverts from the standard use the Kullback-Leibler has in information theory, where a higher divergence means a higher information loss.

Table B3: Effect of cultural proximity on trade, intensive and extensive margins, Kullback-Leibler

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable	Log Sales	Log Transactions	Log Sales	Log Transactions	Trade Indicator	Trade Indicator
KL_{sym}	-0.004***	-0.003**	-0.005***	-0.003**	-0.00004***	-0.00004***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.00000)	(0.00000)
Log dist.	-0.023	-0.065***			0.00007	
	(0.015)	(0.011)			(0.00005)	
Obs.	32,678	32,678	32,843	32,843	5,606,627	5,628,290
Adj. R2	0.415	0.359	0.410	0.356	0.617	0.0106
FE	Seller, buyer	Seller, buyer	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer	Seller, buyer, origin×dest.

Notes: Columns 1, 2, 3 and 4 show the results of estimating a modified version of Equation (2.1). Columns 5 and 6 show the results of estimating a modified version of Equation (2.2). ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Origin-destination fixed effect considers the district of the seller and the buyer. Standard errors two-way clustered at the seller and buyer level. Standard errors in parentheses. A higher Kullback-Leibler divergence means two firms are socially farther away. Number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia, Guimarães, and Zylkin 2019).

Table B4: Effect of cultural proximity on prices, Kullback-Leibler

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable	Log Prices	Log Prices	Log Prices	Log Prices	Log Prices	Log Prices
KL_{sym}	0.003**	0.003**	0.003**	0.002*	0.002**	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log dist.	0.023	0.023	0.028*			
	(0.016)	(0.016)	(0.017)			
Obs.	230,744	230,744	226,645	235,001	236,617	230,900
Adj. R2	0.932	0.932	0.935	0.933	0.925	0.936
FE	Seller×HS, buyer	Seller×HS, buyer, month	Seller×HS, buyer, month×HS	Seller×HS, buyer, origin×dest.	Seller×HS, buyer, month, origin×dest.	Seller×HS, buyer, month×HS, origin×dest.

Notes: This table shows the results of estimating a modified version of Equation (2.3). Good g is defined according to 6-digit HS classification. Prices trimmed by 4-digit HS code at 5 and 95 percent. ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Origin-destination fixed effect considers the district of the seller and the buyer. Standard errors are multi-way clustered at the seller, 4-digit HS and origin-destination level. Standard errors in parentheses. A higher Kullback-Leibler divergence means two firms are socially farther away. Number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia, Guimarães, and Zylkin 2019).

B.3.2 Language

In this section we check if the results we find are driven by language similarity. To do so, we follow the two language similarity measures from **kone2018internal**. Define ϑ_i^l as the share of people with mother tongue l in district i. Then, the common language measure between districts i and j is

$$commlang_{ij} = \sum_l \vartheta_i^l \vartheta_j^l.$$

We can also define a language overlap measure, defined as

$$overlang_{ij} = \sum_l \min \left\{ \vartheta_i^l, \vartheta_j^l \right\}.$$

In both cases, the larger the measures, the less likely it should be for people in these districts to face communication barriers. Table B5 presents the results of the intensive margin regression after considering the language measures. We find that none of the measures is statistically significant. This suggests that the cultural proximity result is not driven by firms sharing the same language.

Table B5: Effect of cultural proximity and language on trade, intensive margin

	(1)	(2)	(3)	(4)
Dep. Variable	Log Sales	Log Transactions	Log Sales	Log Transactions
BC	0.108***	0.068**	0.108***	0.068**
	(0.033)	(0.028)	(0.033)	(0.028)
commlang	-0.322	-0.126		
	(0.389)	(0.305)		
overlang			-0.419	-0.061
			(0.406)	(0.324)
Log dist.	-0.025*	-0.065***	-0.029*	-0.065***
	(0.015)	(0.012)	(0.016)	(0.013)
Obs.	30,703	30,703	30,703	30,703
Adj. R2	0.409	0.357	0.409	0.357
FE	Seller, buyer	Seller, buyer	Seller, buyer	Seller, buyer

Notes: This table shows the results of estimating a modified version of Equation (2.1). ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Standard errors two-way clustered at the seller and buyer level. Standard errors in parentheses. A higher Kullback-Leibler divergence means two firms are socially farther away.

B.3.3 Goods specialization

The cultural groups in India are, in many cases, defined by the production of specific goods (Munshi 2019).[48] In this section we study if the reason behind the cultural proximity results is actually cultural groups specializing in the production of certain goods and, given this, forming special bonds with their specific set of buyers.

First, we assign each firm to a unique cultural group. We do this by assigning each firm to the cultural group for which it has the highest probability of belonging to. In second place, we see which is the most important 4-digit HS code in terms of sales and purchases for each cultural group. We then match each firm to which is the good its cultural group specializes in selling and buying. Working with a version of our dataset at the seller-buyer-good level we run the regression

$$\ln n_g(\nu,\omega,t) = \iota_{\nu\times g} + \iota_{g\times t} + \iota_{\omega} + \delta BC(\nu,\omega) + \xi\left(BC(\nu,\omega)\times \mathbb{I}_g^{spec}\right) + \eta \ln dist(\nu,\omega) + \epsilon_g(\nu,\omega), \tag{B1}$$

where $\mathbb{I}_g^{spec}$ indicates if the good being traded is one in which either the cultural group of the selling firm specializes in selling or the cultural group of the buying firm specializes in buying. Table B6 presents the results for the sales.

First, if the cultural proximity results were only driven by cultural groups producing specific specialized goods, then we would expect the term on cultural proximity to be close to zero, and on the interactions to be statistically different from zero. However, we find that cultural proximity matters for all types of goods: for those in which a cultural group specializes and for those in which a cultural group does not specialize too.

Second, in Column 2 we find that the coefficient on the interaction term is positive and statistically significant. Nevertheless, we lose this statistical significance after controlling for additional variables in Column 4. This could point to cultural proximity mattering more for those goods in which cultural groups specialize in buying, but the result is not conclusive enough.

[48] We can also understand this as certain cultural groups specializing in certain occupations.

Table B6: Effect of cultural proximity on trade by good specialization, intensive margin

	(1)	(2)	(3)	(4)
Dep. Variable	Log Sales	Log Sales	Log Sales	Log Sales
BC	0.072***	0.071***	0.064***	0.064***
	(0.026)	(0.025)	(0.023)	(0.023)
$BC \times \mathbb{I}_g^{spec,seller}$	-0.016		0.135	
	(0.160)		(0.304)	
$BC \times \mathbb{I}_g^{spec,buyer}$		0.152***		0.185
		(0.008)		(0.118)
Obs.	226,039	226,039	229,719	229,719
Adj. R2	0.853	0.853	0.854	0.854
FE	Seller×HS, buyer, month×HS	Seller×HS, buyer, month×HS	Seller×HS, buyer, month×HS, origin×dest.	Seller×HS, buyer, month×HS, origin×dest.

Notes: This table shows the results of estimating Equation (B1). ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Good g is defined according to 6-digit HS classification. Sales trimmed by 4-digit HS code at 5 and 95 percent. Origin-destination fixed effect considers the district of the seller and the buyer. Standard errors two-way clustered at the seller and 4-digit HS level. Standard errors in parentheses. The higher the Bhattacharyya coefficient, the culturally closer two firms are. Number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia, Guimarães, and Zylkin 2019). $\mathbb{I}_g^{spec,seller}$ indicates the good g is the good in which the seller's cultural group specializes in selling. $\mathbb{I}_g^{spec,buyer}$ indicates the good g is the good in which the buyer's cultural group specializes in buying.

B.3.4 Number of varieties sold and bought

In this part we analyze whether firms that the cultural proximity results prevail for firms that sell and buy more varieties of goods. To measure this, we count how many 4-digit HS codes a firm buys or sells. Table B7 presents the results for the intensive margin, following a modified version of Equation 2.1. In our specifications $varieties_{\nu}^{sold}$ and $varieties_{\nu}^{bought}$ refer to the number of varieties sold and bought by the seller, while $varieties_{\omega}^{sold}$ and $varieties_{\omega}^{bought}$ refer to the number of varieties sold and bought by the buyer.

The results point to the effects of cultural proximity on trade being stronger when firms buy and sell more varieties. Our interpretation of these findings is that firms that buy and sell more varieties of goods have to face more contracting frictions, caused by having to negotiate more contracts. Then, these firms, in order to minimize their load of contracting frictions, will rely more on trading with counterparts in which they trust. Moreover, this explanation based on trust is compatible with the results related to differentiated goods from Section 2.3.1. In both cases we posit that the intensity of trade is driven by trust between firms, a coping mechanism to market imperfections in India.

Table B7: Effect of cultural proximity on trade by number of varieties, intensive margin

	(1)	(2)	(3)	(4)
Dep. Variable	Log Sales	Log Sales	Log Sales	Log Sales
BC	0.111***	0.090**	0.107***	0.097**
	(0.040)	(0.040)	(0.035)	(0.039)
$BC \times varieties_{\nu}^{sold}$	0.089			
	(0.126)			
$BC \times varieties_{\nu}^{bought}$		0.121		
		(0.084)		
$BC \times varieties_{\omega}^{sold}$			0.112**	
			(0.051)	
$BC \times varieties_{\omega}^{bought}$				0.068
				(0.043)
Obs.	32,843	32,843	32,843	32,843
Adj. R2	0.410	0.410	0.410	0.410
FE	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.

	(5)	(6)	(7)	(8)
Dep. Variable	Log Transactions	Log Transactions	Log Transactions	Log Transactions
BC	0.056*	0.030	0.056*	0.042
	(0.032)	(0.032)	(0.029)	(0.032)
$BC \times varieties_{\nu}^{sold}$	0.095			
	(0.105)			
$BC \times varieties_{\nu}^{bought}$		0.141**		
		(0.067)		
$BC \times varieties_{\omega}^{sold}$			0.104**	
			(0.042)	
$BC \times varieties_{\omega}^{bought}$				0.071**
				(0.036)
Obs.	32,843	32,843	32,843	32,843
Adj. R2	0.356	0.357	0.357	0.357
FE	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.

Notes: This table shows the results of estimating a modified version of Equation (2.1). ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Origin-destination fixed effect considers the district of the seller and the buyer. Standard errors two-way clustered at the seller and buyer level. Standard errors in parentheses. The higher the Bhattacharyya coefficient, the culturally closer two firms are. $varieties_{\nu}^{sold}$ and $varieties_{\nu}^{bought}$ refer to the number of different HS codes at the 4-digit level sold and bought by the seller divided by 100, respectively. $varieties_{\omega}^{sold}$ and $varieties_{\omega}^{bought}$ refer to the number of different HS codes at the 4-digit level sold and bought by the buyer divided by 100, respectively.

B.4 Model derivations

In this section we include details about the derivations of the theoretical model.

Firms A unique variety ω is produced by a single firm which minimizes its unit cost of production subject to its production technology, so

$$\begin{aligned}\min_{\{m(\nu,\omega)\}} \quad \int_{\nu\in\Omega(\omega)} m\left(\nu,\omega\right)p\left(\nu,\omega\right)d\nu + wl\left(\omega\right),s.t.\\ y\left(\omega\right) = \kappa_{\alpha}z\left(\omega\right)l\left(\omega\right)^{\alpha}m\left(\omega\right)^{1-\alpha},\\ m\left(\omega\right) = \left(\int_{\nu\in\Omega(\omega)} m\left(\nu,\omega\right)^{\frac{\sigma-1}{\sigma}}d\nu\right)^{\frac{\sigma}{\sigma-1}},\\ y\left(\omega\right) = 1.\end{aligned}$$

Merge the first and third constraints, such that

$$\begin{aligned}y\left(\omega\right) &= \kappa_{\alpha}z\left(\omega\right)l\left(\omega\right)^{\alpha}m\left(\omega\right)^{1-\alpha},\\ 1 &= \kappa_{\alpha}z\left(\omega\right)l\left(\omega\right)^{\alpha}m\left(\omega\right)^{1-\alpha},\\ l\left(\omega\right)^{\alpha} &= \frac{1}{\kappa_{\alpha}z\left(\omega\right)m\left(\omega\right)^{1-\alpha}},\\ &= \kappa_{\alpha}^{-1}z\left(\omega\right)^{-1}m\left(\omega\right)^{\alpha-1},\\ l\left(\omega\right) &= \kappa_{\alpha}^{-\frac{1}{\alpha}}z\left(\omega\right)^{-\frac{1}{\alpha}}m\left(\omega\right)^{\frac{\alpha-1}{\alpha}}.\end{aligned}$$

Rewrite the minimization problem, such that

$$\begin{aligned}&\min_{\{m(\nu,\omega)\}} \int_{\nu\in\Omega(\omega)} m\left(\nu,\omega\right)p\left(\nu,\omega\right)d\nu + wl\left(\omega\right),\\ &\int_{\nu\in\Omega(\omega)} m\left(\nu,\omega\right)p\left(\nu,\omega\right)d\nu + \kappa_{\alpha}^{-\frac{1}{\alpha}}wz\left(\omega\right)^{-\frac{1}{\alpha}}m\left(\omega\right)^{\frac{\alpha-1}{\alpha}},\\ &\int_{\nu\in\Omega(\omega)} m\left(\nu,\omega\right)p\left(\nu,\omega\right)d\nu + \kappa_{\alpha}^{-\frac{1}{\alpha}}wz\left(\omega\right)^{-\frac{1}{\alpha}}\left(\int_{\nu\in\Omega(\omega)} m\left(\nu,\omega\right)^{\frac{\sigma-1}{\sigma}}d\nu\right)^{\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}}.\end{aligned}$$

The first order condition with respect to $m(\nu,\omega)$ is

$$
\begin{aligned}
0 &= p(\nu,\omega) + \kappa_\alpha^{-\frac{1}{\alpha}} w z(\omega)^{-\frac{1}{\alpha}} \left(\frac{\sigma}{\sigma-1} \frac{\alpha-1}{\alpha} \right) (\dots)^{\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1} \left(\frac{\sigma-1}{\sigma} \right) m(\nu,\omega)^{\frac{\sigma-1}{\sigma}-1}, \\
p(\nu,\omega) &= \kappa_\alpha^{-\frac{1}{\alpha}} \left(\frac{1-\alpha}{\alpha} \right) w z(\omega)^{-\frac{1}{\alpha}} (\dots)^{\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1} m(\nu,\omega)^{-\frac{1}{\sigma}}, \\
m(\nu,\omega)^{\frac{1}{\sigma}} &= \frac{\kappa_\alpha^{-\frac{1}{\alpha}} \left(\frac{1-\alpha}{\alpha} \right) w z(\omega)^{-\frac{1}{\alpha}} (\dots)^{\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1}}{p(\nu,\omega)}, \\
m(\nu,\omega) &= \frac{\kappa_\alpha^{-\frac{\sigma}{\alpha}} \left(\frac{1-\alpha}{\alpha} \right)^\sigma w^\sigma z(\omega)^{-\frac{\sigma}{\alpha}} (\dots)^{\sigma\left(\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1\right)}}{p(\nu,\omega)^\sigma}.
\end{aligned}
$$

Now, the first order condition with respect to $m(\nu,\omega)$ is

$$
m(\nu,\omega) = \frac{\kappa_\alpha^{-\frac{\sigma}{\alpha}} \left(\frac{1-\alpha}{\alpha} \right)^\sigma w^\sigma z(\omega)^{-\frac{\sigma}{\alpha}} (\dots)^{\sigma\left(\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1\right)}}{p(\nu',\omega)^\sigma}.
$$

We divide both first order conditions, such that

$$
\begin{aligned}
\frac{m(\nu,\omega)}{m(\nu',\omega)} &= \frac{\frac{\kappa_\alpha^{-\frac{\sigma}{\alpha}} \left(\frac{1-\alpha}{\alpha} \right)^\sigma w^\sigma z(\omega)^{-\frac{\sigma}{\alpha}} (\dots)^{\sigma\left(\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1\right)}}{p(\nu,\omega)^\sigma}}{\frac{\kappa_\alpha^{-\frac{\sigma}{\alpha}} \left(\frac{1-\alpha}{\alpha} \right)^\sigma w^\sigma z(\omega)^{-\frac{\sigma}{\alpha}} (\dots)^{\sigma\left(\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1\right)}}{p(\nu',\omega)^\sigma}}, \\
&= \frac{\frac{z(\omega)^{-\frac{\sigma}{\alpha}}}{p(\nu,\omega)^\sigma}}{\frac{z(\omega)^{-\frac{\sigma}{\alpha}}}{p(\nu',\omega)^\sigma}}, \\
&= \frac{p(\nu',\omega)^\sigma}{p(\nu,\omega)^\sigma}, \\
m(\nu',\omega) &= \frac{p(\nu,\omega)^\sigma m(\nu,\omega)}{p(\nu',\omega)^\sigma}.
\end{aligned}
$$

We plug this expression back into the expression for the composite of intermediates, so

$$\begin{aligned}
m(\omega) &= \left(\int_{\nu'\in\Omega(\omega)} m(\nu',\omega)^{\frac{\sigma-1}{\sigma}} d\nu\right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\int_{\nu'\in\Omega(\omega)} \left(\frac{p(\nu,\omega)^{\sigma} m(\nu,\omega)}{p(\nu',\omega)^{\sigma}}\right)^{\frac{\sigma-1}{\sigma}} d\nu\right)^{\frac{\sigma}{\sigma-1}}, \\
&= p(\nu,\omega)^{\sigma} m(\nu,\omega) \underbrace{\left(\int_{\nu'\in\Omega(\omega)} p(\nu',\omega)^{1-\sigma} d\nu\right)^{\frac{\sigma}{\sigma-1}}}_{=\left(P(\omega)^{1-\sigma}\right)^{\frac{\sigma}{\sigma-1}}}, \\
&= p(\nu,\omega)^{\sigma} m(\nu,\omega) \left(P(\omega)^{1-\sigma}\right)^{\frac{\sigma}{\sigma-1}}, \\
&= p(\nu,\omega)^{\sigma} m(\nu,\omega) P(\omega)^{-\sigma}, \\
&= m(\omega) p(\nu,\omega)^{-\sigma} P(\omega)^{\sigma}, \\
p(\nu,\omega) m(\nu,\omega) &= m(\omega) p(\nu,\omega)^{1-\sigma} P(\omega)^{\sigma}, \\
n(\nu,\omega) &= P(\omega) m(\omega) p(\nu,\omega)^{1-\sigma} P(\omega)^{\sigma-1}, \\
&= N(\omega) p(\nu,\omega)^{1-\sigma} P(\omega)^{\sigma-1},
\end{aligned}$$

which is the demand of firm ω from variety ν, where $P(\omega)^{1-\sigma} = \int_{\nu\in\Omega(\omega)} p(\nu,\omega)^{1-\sigma} d\nu$ is the price index faced by firm ω, $n(\nu,\omega) = p(\nu,\omega) m(\nu,\omega)$ is the expenditure of ω on variety ν, and $N(\omega) = P(\omega) m(\omega)$ is the total expenditure of firm ω.

The expression for unit cost of production is

$$\begin{aligned}
c(\omega) &= \frac{w^{\alpha} P(\omega)^{1-\alpha}}{z(\omega)}, \\
&= \frac{P(\omega)^{1-\alpha}}{z(\omega)},
\end{aligned}$$

where wages $w = 1$ is the numeraire price.

Now, firms engage in monopolistic competition since they produce a unique variety. In particular, firm ν maximizes profits by selling its good to buyers ω subject to the demand for its intermediate, so

$$\begin{aligned}
\max_{\{p(\nu,\omega)\}} \quad & \int_{\omega\in\Omega(\nu)} \left(p(\nu,\omega) - d(\nu,\omega) c(\nu)\right) m(\nu,\omega), s.t. \\
& m(\nu,\omega) = m(\omega) p(\nu,\omega)^{-\sigma} P(\omega)^{\sigma},
\end{aligned}$$

where $d(\nu,\omega)$ is the iceberg cost of firm ν selling to ω. Rewrite the profit function $\pi(\nu,\omega)$, such that

$$
\begin{aligned}
\pi\left(\nu,\omega\right) &= \left(p\left(\nu,\omega\right) - d\left(\nu,\omega\right)c\left(\nu\right)\right)m\left(\nu,\omega\right), \\
&= p\left(\nu,\omega\right)m\left(\nu,\omega\right) - d\left(\nu,\omega\right)c\left(\nu\right)m\left(\nu,\omega\right), \\
&= p\left(\nu,\omega\right)m\left(\omega\right)p\left(\nu,\omega\right)^{-\sigma}P\left(\omega\right)^{\sigma} - d\left(\nu,\omega\right)c\left(\nu\right)m\left(\omega\right)p\left(\nu,\omega\right)^{-\sigma}P\left(\omega\right)^{\sigma}, \\
&= m\left(\omega\right)p\left(\nu,\omega\right)^{1-\sigma}P\left(\omega\right)^{\sigma} - d\left(\nu,\omega\right)c\left(\nu\right)m\left(\omega\right)p\left(\nu,\omega\right)^{-\sigma}P\left(\omega\right)^{\sigma}.
\end{aligned}
$$

The first order condition is

$$
\begin{aligned}
\left[p\left(\nu,\omega\right)\right] : &\left(1-\sigma\right)m\left(\omega\right)p\left(\nu,\omega\right)^{-\sigma}P\left(\omega\right)^{\sigma} \\
&- \left(-\sigma\right)d\left(\nu,\omega\right)c\left(\nu\right)m\left(\omega\right)p\left(\nu,\omega\right)^{-\sigma-1}P\left(\omega\right)^{\sigma} = 0, \\
\left(\sigma-1\right)m\left(\omega\right)p\left(\nu,\omega\right)^{-\sigma}P\left(\omega\right)^{\sigma} &= \sigma d\left(\nu,\omega\right)c\left(\nu\right)m\left(\omega\right)p\left(\nu,\omega\right)^{-\sigma-1}P\left(\omega\right)^{\sigma}, \\
\left(\sigma-1\right) &= \sigma d\left(\nu,\omega\right)c\left(\nu\right)p\left(\nu,\omega\right)^{-1}, \\
p\left(\nu,\omega\right) &= \left(\frac{\sigma}{\sigma-1}\right)c\left(\nu\right)d\left(\nu,\omega\right), \\
&= \mu c\left(\nu\right)d\left(\nu,\omega\right),
\end{aligned}
$$

where $\mu = \frac{\sigma}{\sigma-1}$ is the markup.

Households A representative household maximizes its utility subject to its budget constraint, so

$$\max_{\{y(\omega)\}} \quad \left(\int_{\omega\in\Omega} y(\omega)^{\frac{\sigma-1}{\sigma}}\,d\omega\right)^{\frac{\sigma}{\sigma-1}}, \; s.t. \; \int_{\omega\in\Omega} P(\omega)\,y(\omega)\,d\omega \leq Y.$$

The first order condition with respect to firm ω is

$$\begin{aligned}
&[y(\omega)] : \left(\frac{\sigma}{\sigma-1}\right)(\dots)^{\frac{\sigma}{\sigma-1}-1}\left(\frac{\sigma-1}{\sigma}\right) y(\omega)^{\frac{\sigma-1}{\sigma}-1} = \lambda P(\omega),\\
&\lambda P(\omega) = (\dots)^{\frac{\sigma}{\sigma-1}-1}\, y(\omega)^{-\frac{1}{\sigma}},
\end{aligned}$$

where λ is the Lagrangian multiplier of the budget constraint, and $(\dots)$ is an aggregate term we do not write down since it will cancel out during the derivation. Now, the first order condition with respect to another firm ω' is

$$\lambda P(\omega') = (\dots)^{\frac{\sigma}{\sigma-1}-1}\, y(\omega')^{-\frac{1}{\sigma}}.$$

We then divide both first order conditions, such that

$$\begin{aligned}
\frac{\lambda P(\omega)}{\lambda P(\omega')} &= \frac{(\dots)^{\frac{\sigma}{\sigma-1}-1}\, y(\omega)^{-\frac{1}{\sigma}}}{(\dots)^{\frac{\sigma}{\sigma-1}-1}\, y(\omega')^{-\frac{1}{\sigma}}},\\
\frac{P(\omega)}{P(\omega')} &= \frac{y(\omega)^{-\frac{1}{\sigma}}}{y(\omega')^{-\frac{1}{\sigma}}},\\
&= \frac{y(\omega')^{\frac{1}{\sigma}}}{y(\omega)^{\frac{1}{\sigma}}},\\
y(\omega')^{\frac{1}{\sigma}} &= y(\omega)^{\frac{1}{\sigma}}\,\frac{P(\omega)}{P(\omega')},\\
y(\omega') &= y(\omega)\left(\frac{P(\omega)}{P(\omega')}\right)^{\sigma}.
\end{aligned}$$

We plug this demand back in the budget constraint, which holds with equality, so

$$\begin{aligned}
Y &= \int_{\omega' \in \Omega} P(\omega')\, y(\omega')\, d\omega, \\
&= \int_{\omega' \in \Omega} P(\omega') \left[y(\omega) \left(\frac{P(\omega)}{P(\omega')} \right)^{\sigma} \right] d\omega, \\
&= y(\omega)\, P(\omega)^{\sigma} \underbrace{\int_{\omega' \in \Omega} P(\omega')^{1-\sigma}\, d\omega}_{=P^{1-\sigma}}, \\
&= y(\omega)\, P(\omega)^{\sigma}\, P^{1-\sigma}, \\
&= (P(\omega)\, y(\omega))\, P(\omega)^{\sigma-1}\, P^{1-\sigma}, \\
&= x(\omega)\, P(\omega)^{\sigma-1}\, P^{1-\sigma}, \\
x(\omega) &= P(\omega)^{1-\sigma}\, P^{\sigma-1} Y,
\end{aligned}$$

which is the demand function for the unique variety of firm ω, where $P^{1-\sigma} = \int_{\omega \in \Omega} P(\omega)^{1-\sigma}\, d\omega$ is the CES aggregate price index, and $x(\omega) = P(\omega)\, y(\omega)$ is the expenditure on variety ω.

Gravity of intermediates By plugging the pricing equation in the demand of firm ω for intermediates from firm ν, we derive the firm-level gravity equation

$$\begin{aligned}
n(\nu,\omega) &= p(\nu,\omega)^{1-\sigma} P(\omega)^{\sigma-1} N(\omega), \\
&= (\mu c(\nu) d(\nu,\omega))^{1-\sigma} P(\omega)^{\sigma-1} N(\omega), \\
&= \mu^{1-\sigma} d(\nu,\omega)^{1-\sigma} c(\nu)^{1-\sigma} P(\omega)^{\sigma-1} N(\omega), \\
\log(n(\nu,\omega)) &= \log\left(\mu^{1-\sigma} d(\nu,\omega)^{1-\sigma} c(\nu)^{1-\sigma} P(\omega)^{\sigma-1} N(\omega)\right), \\
&= \log\left(\mu^{1-\sigma}\right) + \log\left(c(\nu)^{1-\sigma}\right) + \log\left(P(\omega)^{\sigma-1} N(\omega)\right) + \log\left(d(\nu,\omega)^{1-\sigma}\right), \\
&= \iota + \iota_\nu + \iota_\omega + (1-\sigma)\log(d(\nu,\omega)),
\end{aligned}$$

where ι is an intercept, ι_ν are seller fixed effects, and ι_ω are buyer fixed effects.

Recursive expression for prices. Consider the expression for the CES price index, so

$$
\begin{aligned}
P(\omega)^{1-\sigma} &= \int_{\nu\in\Omega(\omega)} p(\nu,\omega)^{1-\sigma}\, d\nu, \\
P(z')^{1-\sigma} &= \int p(z,z')^{1-\sigma}\, l(z,z')\, dG(z), \\
&= \int \left(\left(\frac{\sigma}{\sigma-1}\right) c(z)\, d(z,z')\right)^{1-\sigma} l(z,z')\, dG(z), \\
&= \mu^{1-\sigma} \int (c(z)\, d(z,z'))^{1-\sigma}\, dG(z), \\
&= \mu^{1-\sigma} \int \left(\frac{P(z)^{1-\alpha}}{z} d(z,z')\right)^{1-\sigma} l(z,z')\, dG(z), \\
&= \mu^{1-\sigma} \int \left(\frac{P(z)^{1-\alpha}}{z} d(z,z')\right)^{1-\sigma} l(z,z')\, dG(z), \\
&= \mu^{1-\sigma} \int P(z)^{(1-\alpha)(1-\sigma)}\, z^{\sigma-1} d(z,z')^{1-\sigma}\, l(z,z')\, dG(z).
\end{aligned}
$$

That is, the price index for firms of productivity z' can be expressed as a function of all other price indexes of firms z. This forms a system of equations we can solve.

Total sales. Consider the expression for total sales (i.e. sales to the household and firms), so

$$
\begin{aligned}
S(\nu) &= x(\nu) + \int_{\omega\in\Omega(\nu)} n(\nu,\omega)\, d\omega, \\
S(z) &= x(z) + \int n(z,z')\, l(z,z')\, dG(z'), \\
&= P(z)^{1-\sigma} P^{\sigma-1} Y \\
&+ \int \left[\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} d(z,z')^{1-\sigma} c(z)^{1-\sigma} P(z')^{\sigma-1} N(z') \right] l(z,z')\, dG(z'), \\
&= P(z)^{1-\sigma} P^{\sigma-1} Y \\
&+ \int \left[\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} d(z,z')^{1-\sigma} \left[\frac{P(z)^{1-\alpha}}{z}\right]^{1-\sigma} P(z')^{\sigma-1} N(z') \right] l(z,z')\, dG(z'), \\
&= P(z)^{1-\sigma} P^{\sigma-1} Y \\
&+ \left[\left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} \left(\frac{P(z)^{1-\alpha}}{z}\right)^{1-\sigma} \right] \int \left[d(z,z')^{1-\sigma} P(z')^{\sigma-1} N(z') \right] l(z,z')\, dG(z'), \\
&= \frac{P(z)^{1-\sigma} Y}{P^{1-\sigma}} \\
&+ \left[\mu^{1-\sigma} P(z)^{(1-\alpha)(1-\sigma)} z^{\sigma-1} \right] \int \left[d(z,z')^{1-\sigma} P(z')^{\sigma-1} N(z') \right] l(z,z')\, dG(z'), \\
&= \frac{\left[\left(\frac{\sigma}{\sigma-1}\right) c(z) D(z)\right]^{1-\sigma} Y}{P^{1-\sigma}} \\
&+ \left[\mu^{1-\sigma} P(z)^{(1-\alpha)(1-\sigma)} z^{\sigma-1} \right] \int \left[d(z,z')^{1-\sigma} P(z')^{\sigma-1} \left(\frac{(1-\alpha) S(z')}{\mu}\right) \right] l(z,z')\, dG(z'), \\
&= \left(\mu^{1-\sigma} P(z)^{(1-\alpha)(1-\sigma)} z^{\sigma-1} D(z)^{1-\sigma} \right) \frac{Y}{P^{1-\sigma}} \\
&+ \left[\mu^{1-\sigma} P(z)^{(1-\alpha)(1-\sigma)} z^{\sigma-1} \right] \left[\frac{1-\alpha}{\mu}\right] \int \left[d(z,z')^{1-\sigma} P(z')^{\sigma-1} S(z') \right] l(z,z')\, dG(z'), \\
&= \left[\mu^{1-\sigma} P(\nu)^{(1-\alpha)(1-\sigma)} z^{\sigma-1} \right] \\
&\left[\frac{Y}{P^{1-\sigma}} D(z)^{1-\sigma} + \left(\frac{1-\alpha}{\mu}\right) \left(\int \left[d(z,z')^{1-\sigma} P(z')^{\sigma-1} S(z') \right] l(z,z')\, dG(z') \right) \right],
\end{aligned}
$$

where we use the fact that $N(z') = \frac{(1-\alpha)S(z')}{\mu}$. Given prices $P(z)$, this forms a system of equations for sales we can solve.

Appendices to Chapter III

C.1 Additional tables

Table C1: 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 of two panels. Panel A contains information about the 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 for 2020.

Table C2: Distribution of economic activity by industry and type of transaction

HS section	Sales share	Purchase share
Animals	1.5034	.7723
Vegetables	15.2982	11.2945
Fats	2.2934	2.6251
Processed foods	4.2172	5.5548
Minerals	13.1241	10.2353
Chemicals	9.8288	9.0791
Plastics	13.1516	9.1410
Leather	.1618	.1677
Wood	2.5110	1.2130
Wood derivatives	1.0783	1.3598
Textiles	3.6342	6.4576
Clothing	1.3428	.9107
Handicrafts	1.0190	1.9337
Jewelry	1.7005	1.4980
Metal	10.4473	12.1969
Machinery	10.9909	13.5771
Transport equipment	4.7124	8.4147
Surgical instrum.	1.4478	1.6478
Arms and ammo	.0057	.0095
Miscellaneous	1.2263	1.4936
Art	.3043	.4166
Type of transaction		
Within-state	72.6822	52.2224
Inter-state	23.2183	44.5151
Foreign	4.0994	3.2623

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

Table C3: 2SLS, firm-level elasticity of substitution across (at least two) suppliers

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

Notes: IV-2SLS estimates from Equation (3.9). The set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$. That is, a supplier s of buyer b is considered *common* if they also traded during the previous month. We only consider the cases when a buyer traded with at least two common suppliers in a given period. The first stage uses either bilateral or seller-level instruments, as pointed out by rows six and seven. Bilateral instruments correspond to Equation (3.12), while seller-level instruments correspond to Equation (3.11). The first row reports the estimates associated with changes in relative unit values in logs. Standard errors are two-way clustered at the origin and destination state level, and are reported in parentheses below each estimate. The fourth row reports the Kleibergen-Paap F statistic from the first stage. The fifth row reports the implied value for \epsilon, which is 1 minus the estimate on the first row. A product category is a 4-digit HS codes and the treatment period is March-May 2020. The table contains four columns. Each column corresponds to different combinations of instruments and of fixed effects, as pointed out by the last six rows. All specifications include the controls mentioned in the paper.

Table C4: Robustness: 2SLS, firm-level elasticity of substitution across suppliers

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

Notes: IV-2SLS estimates from Equation (3.9). The set of common suppliers of buyer b is $\Omega^*_{i,bj,t} = \Omega_{i,bj,t} \cap \Omega_{i,bj,t-1}$. That is, a supplier s of buyer b is considered *common* if they also traded during the previous month. In all specifications, the first stage uses both bilateral and seller-level instruments as pointed in rows seven and eight. Bilateral instruments correspond to Equation (3.12), while seller-level instruments correspond to Equation (3.11). The first row reports the estimates associated with changes in relative unit values in logs. Standard errors are two-way clustered at the origin and destination state level, and are reported in parentheses below each estimate. The fourth row reports the Kleibergen-Paap F statistic from the first stage. The fifth row reports the implied value for ϵ, which is 1 minus the estimate on the first row. A product category is either 6-digit or 8-digit HS codes as pointed out by the sixth row, and the treatment period is March-May 2020. The table contains four columns. Each column corresponds to different combinations of HS codes and of fixed effects, as pointed out by the last five rows. All specifications include the controls mentioned in the paper.

Table C5: Firm-level elasticities of substitution across suppliers, by HS section

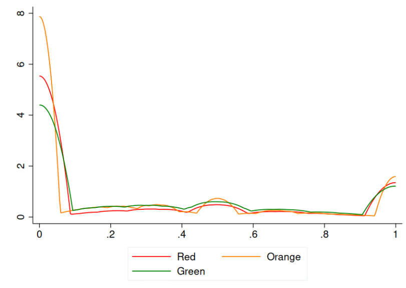
Section	Name	OLS elast.	2SLS elast.
1	Animals	.	.
2	Vegetables	.8082675	.
3	Fats	.825045	.3875365
4	Processed foods	.7458998	1.141
5	Minerals	.8220726	.5755809
6	Chemicals	.	.
7	Plastics	.8097205	.
8	Leather	.	.
9	Wood	.8905362	.
10	Wood derivatives	.8832779	.8700905
11	Textiles	.8635682	1.636
12	Clothing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

Notes: Each row corresponds to an industry, which is defined as a HS section. The second column contains the name of the industry. The third and fourth columns report the estimated elasticities by OLS and 2SLS as in Equation (3.9). Both OLS and 2SLS estimators include HS/time, buyer/time, buyer/HS, and seller/HS fixed effects. Standard errors are two-way clustered at both origin and destination states. All specifications include the controls mentioned in the paper. Elasticities were not reported if there was low statistical power or a weak first stage.

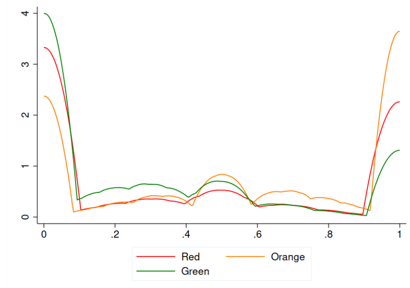
C.2 Additional figures

Figure C1: Distribution of links and sales across lockdown zones

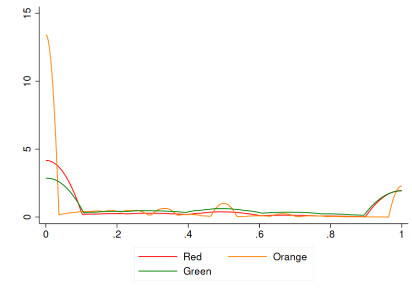
(a) Sellers in Red



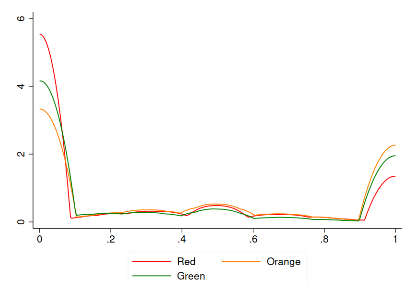
(b) Sellers in Orange



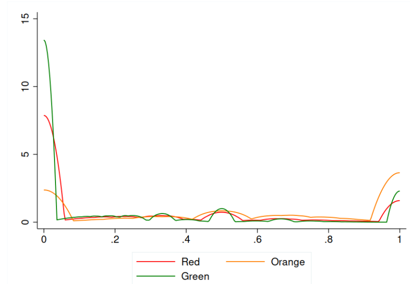
(c) Sellers in Green



(d) Buyers in Red



(e) Buyers in Orange



(f) Buyers in Green

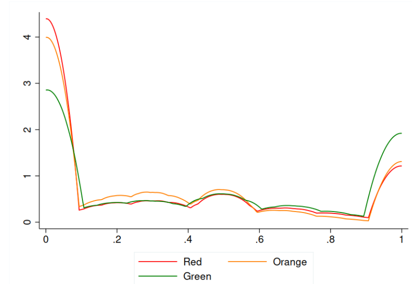
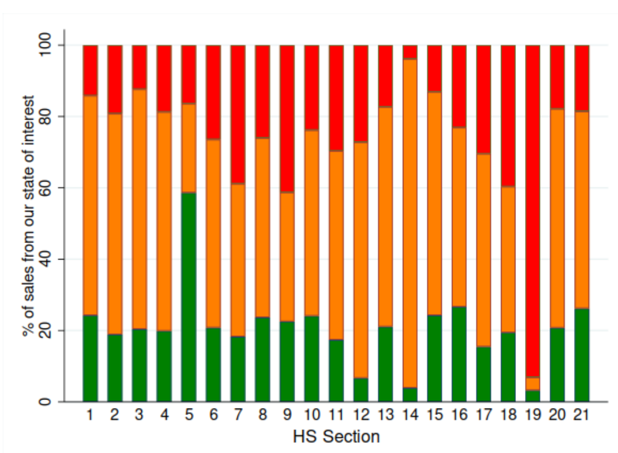
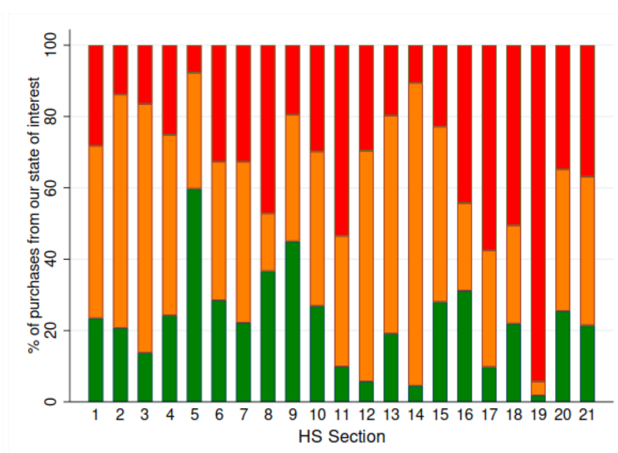


Figure C2: % of sales/purchases, by color of destination districts



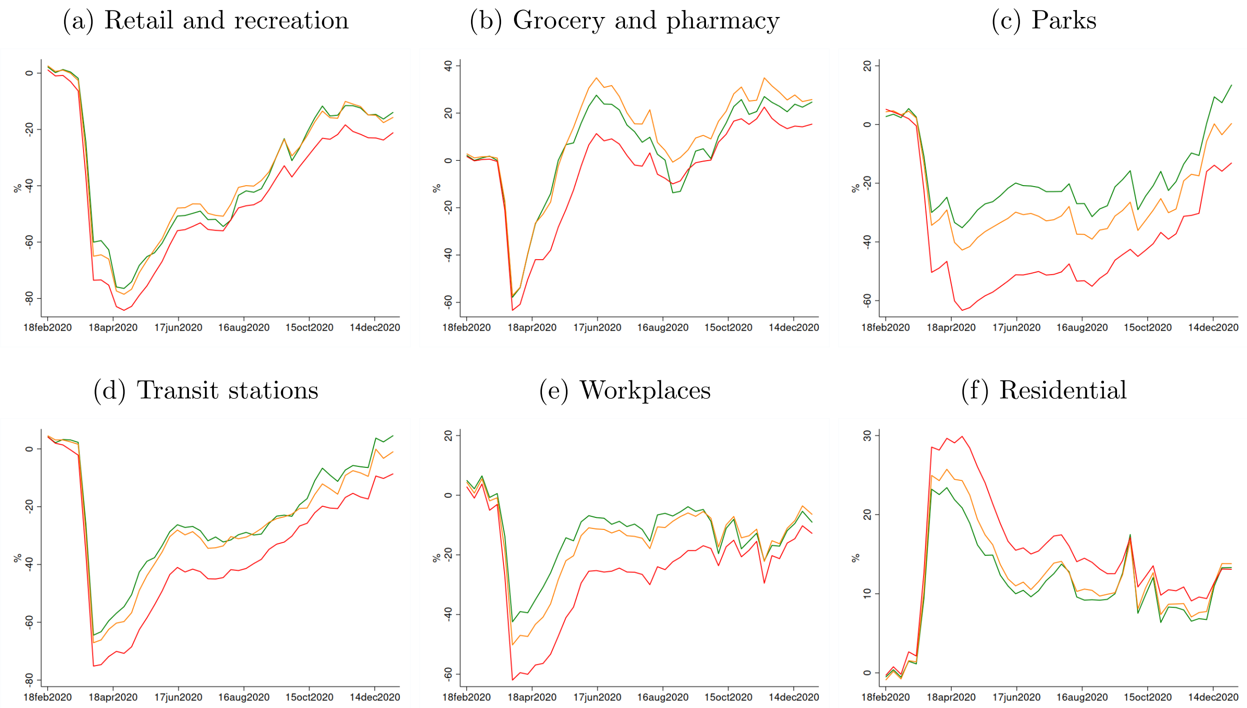
(a) Sales



(b) Purchases

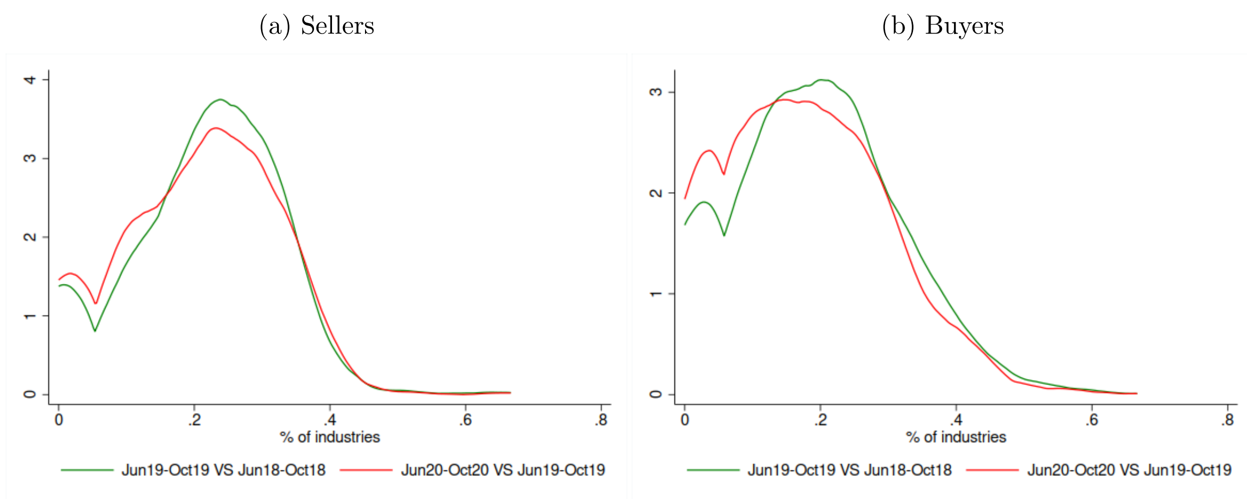
Notes: This figure is comprised by two set of panels. The first six figures are the first panel, and the last two figures are the second panel. First we explain the first panel. In the three upper figures, each panel plots the distribution of the share of buyers located in *Red*, *Orange*, or *Green*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

Figure C3: Google mobility trends by lockdown zone



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 C4: Change in Product Category Links, Before/After Lockdowns



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

C.3 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

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

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 OLS estimates, and use them to predict labor and final sales for all firms.

We scraped data on number of employees and total sales from the website *IndiaMART*, India's largest B2B digital platform. We scraped around 300,000-400,000 firm profiles, and then sent them to the tax authority to be matched with our firm-to-firm trade dataset. The matching 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.[49] 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 then run the following OLS regressions:

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

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

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

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

First, we obtain the centroid of each district in India. Then , we construct a network structure according to the set of centroids. There are many ways to construct a network, so we need to take a stance on how to form the connections between centroids. For each centroid, we generate connections to the k closest centroids according to Euclidean distances.[50] We

[49] 100,000 rupees = 1 Lakh; and 10,000,000 rupees = 1 Crore.

[50] 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

follow Fajgelbaum and Schaal 2020 and consider $k = 8$ such that we consider the main cardinal directions (i.e. north, south, east, west, north-east, south-east, north-west, south-west).

We now run the Dijkstra algorithm. For all district pairs, the algorithm provides us with the list of all districts that comprise the route between the district pair, and the distance of each leg that comprise the route. Using the 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.4 Derivations

C.4.1 Estimation of firm-level elasticities of substitution across suppliers

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

Expression to estimate firm-level elasticities of substitution across suppliers. A firm b selling product $j \in F$ maximizes profits subject to its technology and to a CES bundle of

from one node to another. $k = K$ is a fully-connected network, where all nodes are connected with each other. Running a Dijkstra algorithm on this scenario is trivial since the shortest distance between any pair of nodes is their connection itself. Therefore, a feasible number of connections per node must be $k \in (0, K)$.

intermediate inputs:

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

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

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

where $(\ldots)$ are components that we do not write in detail since they cancel out eventually. 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 $PM_{i,bj} \equiv \sum_{s'} PM_{s'i,bj}$.

Constructing price indices. In this section we derive the expressions that allows us to construct price indexes based on observable data. First, go back to the previous derivation, 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 introduce a time dimension such that

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

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

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

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

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

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

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

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

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

This expression in changes is

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

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

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

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

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

Then, the last term of our expression is

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

So our final expression boils down to

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

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

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

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

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

In changes, the expression for expenditure shares is

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

Then, the geometric mean for expenditure shares is

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

So the final expression we take to the data is

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

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

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

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

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

In logarithms, we have

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

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

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

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

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

If we express this in changes, we get

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

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

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

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 C.3. Combining the expression for changes in travel time due to the lockdown and trade costs, we get the following expression for our seller/buyer level excluded instruments

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

C.4.2 Estimation of firm-level elasticities of substitution across products

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

Expressions to estimate firm-level elasticities of substitution across products. We rewrite the initial maximization problem, so

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

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

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

where $(\ldots)$ are components that we do not write explicitly since they eventually cancel out. 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} \left(1-w_{bl}\right) 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} \left(1-w_{bl}\right) 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 also assume that the overall importance of the composite intermediates is time-invariant, so

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

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

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

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

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

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

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

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

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

7. We then construct unit values by

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

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

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

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

C.5 Simulations using quantitative model

C.5.1 Deriving expression for shock propagation through GDP

In this section, we discuss details of the simulation using the quantitative model. In order to do that, we first recall the different notations used in the paper. N is the number of firms, I is the number of product categories. λ_k is the *Domar* weight of firm or sector k. θ_k is the elasticity of substitution corresponding to the k^{th} reproducible sector. Ω_{li} is the $(l,i)^{th}$ element of the $(N+I+2)$ input output matrix Ω. It therefore measures the direct reliance of l on i as a supplier. ψ_{li} corresponds to the $(l,i)^{th}$ element of the $(N+I+2)$ Leontief inverse, and captures the direct and indirect reliance of l on i as a supplier. The aggregate change in GDP ($\Delta logy$) in response to changes in productivity of firm j ($\delta logA_j$) up to a second order is given by the following:

$$
\begin{aligned}
logy =& \sum_{j=1}^{N} \frac{\partial logy}{\partial logA_j}(\Delta logA_j) \\
&+ \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1,i\neq j}^{N} \frac{\partial^2 logy}{\partial logA_i \partial logA_j}(\Delta logA_i)(\Delta logA_j) + \frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2 logy}{\partial logA_i^2}(\Delta logA_i)^2. \qquad \text{(C1)}
\end{aligned}
$$

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

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

We will now write down the expressions for B and C in matrix form in order to evaluate the second order effects. $\times$ denotes matrix multiplication and $\cdot$ denotes element by element matrix operations.

Quantifying B. To quantify B, the term that mainly captures the second order effects on GDP that operates through changes in firm i's *Domar* weight in response to productivity shocks to firm j, where $j \in N, j \neq i$, we introduce the following intermediate matrices which we will define below: M, N, $Covar1$, $Covar21$, $Covar22$, and $Covar2$. $J_{m,n}$ denotes a matrix of ones of size m by n.

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

Quantifying C. The term C, mainly captures the second order effects on GDP that operates through changes in firm i's *Domar* weight in response to productivity shocks to firm i itself.

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

Matrix form. We can rewrite Equation (C1) as:

$$\begin{aligned}
\Delta logy = \lambda \times \Delta logA + .5 \Big((\theta - 1) \cdot \lambda \Big) \times \Big(Covar1 - Covar2 \Big) + \\
.5 \Bigg(\Big((\theta - 1) \cdot \lambda \Big) \times \Big(\Omega \times (\psi \cdot \psi) - (\Omega \times \psi) \cdot (\Omega \times \psi) \Big) \Bigg) \times \Bigg(\Delta logA \cdot \Delta logA \Bigg)
\end{aligned} \tag{C2}$$

C.5.2 Numerical implementation in Python

Numerically implementing this exercise is challenging due to the sheer size of the firm-to-firm trade network. We have data on 93260 firms across 1293 product categories. This generates a 94,555 by 94,555 input output matrix. The elements inside the input-output matrix are very small as the fraction of a product's output going to a single firm is very small and each product category in turn sources from a large number suppliers. Therefore, and to keep the calculations as precise as possible, we had to use *float64* variable types with these matrices, which resulted in matrices larger than most servers' memories. For instance, the Leontief inverse matrix alone took more than 66GB of storage/memory size. A lot of the calculation's steps required performing matrix multiplication operations on these large matrices. Matrix multiplication is one of the most demanding operations in terms of computing resources in the world of computer science. We break down this computation via a number of state-of-the-art big data computing techniques, thus achieving scalability when applying our techniques to arbitrarily large input output matrices. As detailed firm-to-firm transactions data are becoming more widely available, these techniques will advance the literature quantifying the propagation of shocks through firm networks.

First, we are able to fit datasets larger than RAM using Dask which provides multi-core and distributed and parallel execution on *larger-than-memory* datasets.[51] We use Dask distributed capabilities to add parallelism to the calculations in computing second order effects which require few matrix multiplication operations on large 94,555 by 94,555 matrices.

Second, we use a computer powered with multiple GPUs. GPUs are essential for the numerous matrix multiplications this process involves. To demonstrate this in numbers, computing 10 columns of Leontief inverse matrix (only 0.000001%) takes about 4 days on a powerful server with multiple CPUs, 500GB of RAM and 16 cores. Computing the entire Leontief inverse on a server powered with 4 GPUs took about 1 hour. The part of our work of computing the second order effect, which involves 3 operations of large matrix multiplication would not be practical using CPUs only.

Third, we use the properties of sparse matrices to define matrix multiplications that can ignore large contiguous chunks of zeros, a typical feature of input output matrices.

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

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