The Aggregate Implications of Cultural Proximity*

Brian Cevallos Fujiy
U.S. Census Bureau

Gaurav Khanna†
University of California, San Diego

Hiroshi Toma
University of Michigan

First draft: April, 2022
This draft: November, 2023

Abstract

Emerging economies often feature low-quality institutions, generating micro-level trade frictions. In these settings, firms may rely on cultural-proximity-based informal institutions to overcome such frictions. We quantify the aggregate effects of cultural proximity in a production network. Using new microdata on firm-to-firm trade from India with information on prices, transactions, and caste and religious connections, we find that higher cultural proximity reduces prices and fosters trade at intensive and extensive margins. Our evidence suggests these results are driven by firms trying to overcome frictions imposed by low-quality institutions. Guided by these facts, we propose a quantitative firm-level production network model, where cultural proximity and institutional quality influence trade and matching costs. Our counterfactual exercises indicate that an economy composed of culturally closer firms features lower costs, lower prices, higher sales and higher welfare with respect to an economy with culturally distant firms.

Keywords: cultural proximity, production networks, firm-to-firm trade

JEL Codes: D51, F19, O17

*Previous title: "*Cultural Proximity and Production Networks.*" We thank Treb Allen, Mathias Buehler, Paula Bustos, Vasco Carvalho, Thomas Chaney, Ornella Darova, Kevin Donovan, Ben Faber, Farid Farrokhi, Alan Griffith, Amit Khandelwal, Andrei Levchenko, Kanika Mahajan, Rocco Macchiavello, Alexey Makarin, Yuhei Miyauchi, Joan Monras, Swapnika Rachapalli, Kim Ruhl, Jagadeesh Sivadasan, Sebastian Sotelo, Tillmann von Carnap, Daniel Xu, and Yingyan Zhao for insightful comments. We thank seminar participants at the University of Michigan, BSE Summer Forum, JADE-CEPR-TIME Conference, LMU Economics of Firms and Labor Workshop, MWIEDC Conference, NEUDC Conference, PacDev Conference, STEG Theme Workshop, and WAITS Conference for their useful suggestions. We thank Manaswini Bhalla, Manisha Goel, and Manaswini Rao for sharing their datasets. Eli Mogel and Tyler Spencer provided excellent research assistance. We also thank the 2021 Hellman Fellowship for their generous support for this project. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau.
Fujiy: brian.j.cevallos@census.gov, Khanna: gakhanna@ucsd.edu, Toma: htoma@umich.edu

†Corresponding author.

1 Introduction

Low-quality institutions are prevalent in emerging economies, distorting firm-level production and sourcing decisions. For instance, in production networks, buyers can hold up payments to suppliers when there are low levels of contract enforcement. In the presence of such frictions, agents may rely on cultural-proximity-based informal institutions.

Cultural ties may ease contracting frictions by lowering trade costs and replacing contract enforcement with trust. Yet, culture-based contracts could, in contrast, lead to inefficient trades, as potentially low-productivity but culturally-close firms start trading. These ambiguities in firm-level trade are particularly amplified when considering the aggregate implications for the overall economy. In this paper, we first examine whether trade that is reliant on cultural proximity eases contracting frictions, and then quantify the economy-wide consequences of culture-based trade.

Cultural proximity—shared values, community, religion, among others—is an important informal institution, affecting entrepreneurship (Goraya, 2023), production processes (Ghosh, 2023), industrial investment (Gupta, 2014), human capital and political economy (Bazzi et al., 2020), social behavior (Rao, 2019), conflict (Mitra and Ray, 2014), social mobility (Hnatkovska et al., 2012, 2013), coordination failures (Afridi et al., 2020), competitions (Lowe, 2021), and labor markets (Munshi and Rosenzweig, 2016). On the one hand, cultural ties can help overcome contracting frictions, but on the other, can lead to discrimination and social exclusion, generating a misallocation of resources. This previous work provides limited evidence on the aggregate implications of micro-level culture-based trade.[1]

Quantifying the aggregate implications of cultural proximity on firm-to-firm trade has been challenging for two reasons. First, so far, there has been little data on how firms relate culturally, their trade patterns, and, importantly, at what prices they trade. Second, there are few quantitative frameworks to measure the aggregate consequences of cultural proximity.

We address these challenges in three steps. First, 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. India has a society that follows the parameters of a caste system, which intertwines with the different religious groups. In our context, cultural proximity naturally arises as a product of the inherent hierarchical structure of India's caste system and the different religions. Second, using this

[1]Other work with aggregate consequences focuses on talent misallocation (Hnatkovska et al., 2021) and political outcomes (Bazzi et al., 2020).

unique dataset, we document empirical micro-level patterns of how cultural proximity affects firm-to-firm trade. We find that larger cultural proximity between firms reduces prices and fosters trade at both the intensive and extensive margins. Third, armed with these micro-level patterns, we derive and discipline a theoretical model that quantifies the aggregate effects of cultural proximity. We find that cultural proximity has an aggregate effect: moving from a culturally distant economy to a culturally close economy improves welfare, increases sales, and reduces prices through a trade cost and matching cost channel. However, average productivity falls, as falling costs mean that less productive firms will trade more.

We report three new stylized facts that relate firm-to-firm trade with cultural proximity. 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. That is, the higher the cultural proximity, the higher the trade on the extensive margin as well.

Third, and most importantly, firms that are culturally closer report lower unit prices in their transactions. This is a key empirical result that allows us to conclude that cultural proximity alleviates distortions that lower equilibrium prices. All our 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).

Importantly, we find evidence suggesting that firms rely on cultural proximity to overcome the frictions to trade imposed by low-quality institutions. First, we show that the effects of cultural proximity on trade is driven by differentiated products, which often rely on either formal or informal contract enforcement (Boehm and Oberfield, 2020; Nunn, 2007; Rauch, 1999).[2] Second, cultural proximity matters the most when institutional quality, proxied by the court quality in the districts where the trade partners belong, is particularly low. We argue that, in a setting with low institutional quality and poor contract enforcement, firms that trade differentiated goods rely on informal institutions (i.e. cultural proximity) as a substitute for the imperfect formal ones. We understand these findings as evidence that cultural proximity relates to contract enforcement and trust (Munshi, 2019, 2014).[3]

[2] 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.

[3] We also study additional empirical specifications, all of which point in the same direction: cultural proximity increases trade. We find the effect of cultural proximity on trade is more relevant when firms are exposed to low-quality institutions. The additional specifications also suggest preference-based discrimination and specialization are not behind the empirical results.

We then build a quantitative model of firm-to-firm trade and cultural proximity. Based on our empirical results that show cultural proximity is relevant when institutions are imperfect, we introduce cultural proximity in the model through contract enforcement-based and matching-based microfoundations.

In the model, firms engage in monopolistic competition and produce goods by combining labor and intermediate inputs. Firms sell their goods to a household as final goods and to other firms as intermediates, where sellers endogenously choose buyers. A seller matches with a buyer whenever the profits of doing so are larger than the fixed costs of matching. Guided by our extensive margin result, cultural proximity between sellers and buyers reduces the cost of matching, since culture encodes information that is useful for sellers to decide which buyer to sell to.

Furthermore, guided by our intensive margin results and the role of contract enforcement for differentiated goods, sellers charge a premium to buyers arising from contracting frictions. Given the risk of reneging on the contract or delaying payment, sellers determine the charged premium to buyers, and so affect the intensive margin of trade (Boehm, 2015). Modeling the contracting friction premium and fixed costs of matching as a function of cultural proximity allows us to match our empirical findings, and are a simple way to quantify the aggregate implications of cultural proximity.

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, and that of matching costs to cultural proximity. In line with our findings, the magnitude of these parameters depends on institutional quality. Our model allows us to estimate these parameters externally. Our estimates imply that 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. Importantly, the higher the institutional quality, the less sensitive these parameters become to cultural proximity.

We then quantify the aggregate effects of cultural proximity by studying several counterfactual exercises. First, we quantify the aggregate effect of going from a culturally distant economy to a culturally close economy, by lowering trade and matching costs. We find that welfare and sales increase by 11 percent, while prices fall by 10 percent. We also find that average productivity falls by 0.22 percent, as lower-productivity firms start selling more. This last result may speak of a possible drawback introduced by cultural proximity: there are unproductive firms trading. Indeed, relying on culture-based trade may induce more trade with less productive firms.

In a second counterfactual, we analyze a case where we start with an economy where firms are culturally the closest and move to the baseline case, where cultural proximity is disciplined by the data. This counterfactual shows what the observed distorted economy wins or loses in the aggregate with respect to a frictionless economy. We find welfare, sales, and average matches decrease substantially. Prices rise by 6.61 percent. Lastly, average productivity goes up as unproductive firms lose more in sales than the more productive firms.

Our third counterfactual directly relates to the policy consequences of improving institutional quality. It builds on our empirical court-quality finding and quantifies the aggregate effect of a policy that improves courts. Within the model, this means that the trade and matching costs become less sensitive to cultural proximity. We find that improving courts increases welfare by 7.71 percent, sales by 9.31 percent, and average matches by 0.05 percentage points. Prices go down by 7.15 percent. Average productivity falls, driven by more unproductive firms being able to sell more; once again highlighting an important drawback that culturally-reliant trade may induce more trade by less productive firms.

We contribute to two strands of the literature. First, we speak to the role of cultural proximity on trade (Bandyopadhyay et al., 2008; Guiso et al., 2009; Macchiavello and Morjaria, 2015; Rauch, 1996; Rauch and Trindade, 2002; Schoar et al., 2008; Startz, 2016). Yet, much of this work is on cultural differences in international trade or across administrative regions, which also interplay with non-cultural barriers. In contrast, we use transaction-level firm-to-firm data to explore the effect of culture on trade that does not rely on cross-border variation. This allows us to isolate the mechanisms at a granular level, and to provide a theoretical framework that quantifies the role of cultural proximity in production networks.[4]

Second, we contribute to quantitative work on production networks (Antras et al., 2017; Bernard et al., 2022; Dhyne et al., 2021; Eaton et al., 2022; Huneeus, 2018; Lim, 2018; Oberfield, 2018; Taschereau-Dumouchel, 2019) by embedding culture into a network framework, and showing how cultural proximity shapes inter-firm trade and aggregate welfare. The uniqueness of our data in terms of measuring firm-to-firm transactions and the cultural

[4] In ongoing work, Boken et al. (2022) also shows the role of cultural proximity for inter-firm trade. We mainly distinguish ourselves by leveraging data on prices, which allows us to estimate how cultural proximity influences inter-firm trade through the alleviation of contracting frictions. Additionally, we explore a rich set of mechanisms, as our data contains information to test how cultural proximity matters for inter-firm trade through discrimination (caste hierarchies), order cancellations, firm survival over age, and how cultural proximity matters for firm-level complexity. Leveraging other data, we implement a Heckman selection bias correction model to estimate trade elasticities following Helpman et al. (2008).

group of owners, in combination with substantial variation across cultural groups, allows 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 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 3, we report our stylized facts. In Section 4, we describe the model. In Section 5, we explain how we estimate the key parameters of the model. In Section 6, we analyze counterfactual scenarios. Section 7 concludes.

2 Background, data and construction of variables

To quantify the aggregate effects of cultural proximity, we require a setting where both notions of cultural proximity, and low-quality institutions coexist. The Indian context is ideal as it provides us with this exact setup. In this section, we first discuss the institutional setting in India, where caste and religious divides generate notions of cultural proximity. We then discuss court quality and contract enforcement, before describing our new micro-level dataset that combines firm-to-firm trade with cultural group information.

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 affects the economy. 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, and traders, among others. Finally, the Shudras are the most discriminated against and are the caste formed to be the labor class. Below these groups in the socio-economic hierarchy, were marginalized groups called Dalits.

Varnas are comprised of sub-groups called *Jatis*, determined by factors such as occupation, geography, tribes, or language, and so using Jatis as castes is appropriate for studying economic networks (Munshi, 2019). We also consider religious groups to define other cultural groups. The caste system is inherently based on the Hindu religion, the predominant religion in India. While there are other religions in India that did not historically follow the caste system, they do relate to it today: the other non-Hindu religions work as cultural groups of their own. We leverage information on firm owners belonging to both caste and non-Hindu religious groups to construct our measure of cultural proximity.

2.2 Institutions and Courts in India

International comparison shows that India lags in formal institutional quality. According to the World Governance Indicators (Kaufman and Kraay, 2023) in 2019, India ranked in the lower half worldwide for the rule of law and for control of corruption. In comparison, the United States ranked near the top decile for both indicators.

We focus on court quality. Evidence shows that courts in India are characterized by several delays. Moreover, low-quality courts are pervasive at all levels. The India Justice Report (2023) shows that the total number of cases pending a decision for longer than 10 years at both state and district levels surpassed 52 million in 2023, with the chief reasons behind the long delays being underfunding and understaffing (Rao, 2023).

Court delays mean low contract enforcement, affecting firm-to-firm trade. Anecdotal evidence shows solving commercial contract-related disputes between small and medium-sized firms in a sample of Indian courts can take between 600 and 1400 days, depending on the state.[5] This shows how institutions can be relevant for firms when taking their production and trade decisions, as their low-quality can introduce significant frictions to the functioning of the markets.

2.3 Data

Firm-to-firm trade. We obtain new firm-to-firm trade data for a large Indian state provided by the state's corresponding tax authority.[6] We use daily transaction-level 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 to significantly increase tax compliance.[7]

[5] https://subnational.doingbusiness.org/en/data/exploretopics/enforcing-contracts/india

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

[7] Tax evasion rates are thought to have fallen with the E-Way bill system, given how it is implemented. A selling establishment must online register the transaction, and print out a receipt that the driver of

The state has a diversified production structure, roughly 50 percent 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 double the population of Chile. In addition, we can uniquely measure product-specific prices for each transaction, along with the usual measures of the total value traded.

Each transaction reports a unique tax code identifier for both the seller and buyer. 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 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 (600 USD) must generate E-way bills. Transactions that have values lower than 600 USD can also be registered, but it is not mandatory. There are three types of recorded transactions: (i) within-state trade, (ii) across-state trade, and (iii) international trade. Given our research questions, we ignore international trade.

Firm owner names. Information about the name of firm owners comes from two different sources. The first 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 state. Among these variables, we are provided with the name of the owner, directors and/or 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

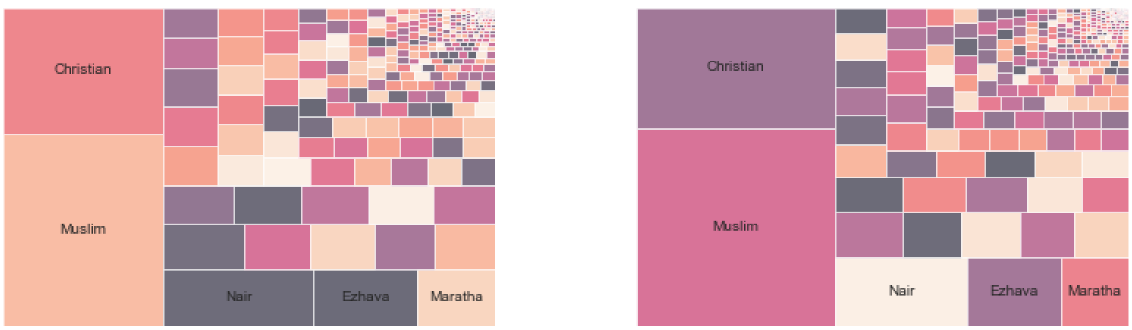
the transportation (usually a truck) must carry with them while transporting the product. If the driver is stopped or checked at any of the numerous checkpoints, and fails to produce a receipt, the goods are confiscated. Furthermore, the earlier VAT regime was only for large firms, and the new GST system was aimed at including smaller firms too. For more details about the new E-Way bill system, see https://docs.ewaybillgst.gov.in/

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

the owners are not Hindu). Their procedure consists of using scraped data from Indian matrimonial websites that contain information on names, castes, and religions. They train a sorting algorithm that uses names as inputs and gives a probability distribution across cultural groups per name as outputs. We match these probability distributions to each owner's 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.

Figure 1: 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 the rectangles reflects the share of sales and purchases.

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 are the same parent firm. Each firm is linked to a unique pincode. Finally, we assign a sector to each firm based on the ISIC 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 1.

2.4 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 $\times$ buyer pair, we construct (i) total sales, (ii) sales at the 6-digit HS good level, (iii) total number of transactions, and (iv) 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. To obtain prices, we calculate the unit values. 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 for 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 $\boldsymbol{\rho_\nu} = [\rho_\nu(1), \ldots, \rho_\nu(X)]$, 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 coefficient (Bhattacharyya, 1943).

The Bhattacharyya coefficient between seller ν and buyer ω measures the level of overlap 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, $BC(\nu, \omega) = 0$, implies the seller has a completely different probability distribution from the buyer's. In our context, this means the seller and the buyer have almost no chance of belonging to the same cultural group or that their cultural proximity is the farthest. On the other hand, $BC(\nu, \omega) = 1$, implies the seller has exactly the same probability distribution as the buyer. This implies that the seller has the same probability of belonging to a group of

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

certain cultural groups as the buyer or that their cultural proximity is the closest possible.[10] In robustness checks, we use the Kullback and Leibler (1951) divergence measure (Appendix C.1). All our results are qualitatively similar, and statistically significant when doing so.

3 Stylized facts

In this section, we show stylized facts about how cultural proximity influences firm-to-firm trade. The importance of this section is twofold. First, we document novel facts about firm-to-firm trade under a setting where low-quality institutions and cultural proximity are prevalent. Second, these facts will be useful to guide our theoretical model in the quantification of the aggregate effects of cultural proximity.

Fact 1: Cultural proximity increases trade at the intensive margin. Figure 2 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) relates to a higher amount of sales and transactions.

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

$$\ln y\left(\nu,\omega\right)=\iota_{\nu}+\iota_{\omega}+\iota_{od}\left(\nu,\omega\right)+\delta BC\left(\nu,\omega\right)+\varepsilon\left(\nu,\omega\right), \tag{1}$$

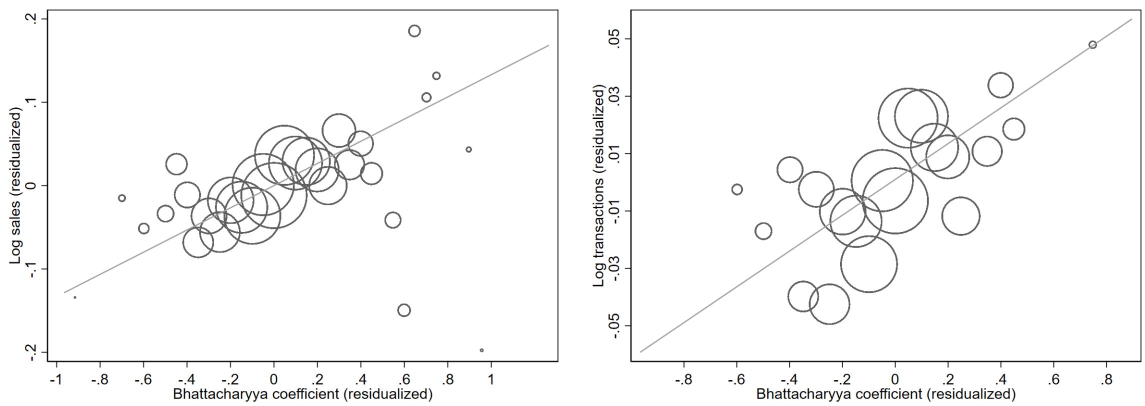
where $y\left(\nu,\omega\right)$ is either the total sales $n\left(\nu,\omega\right)$ or total transactions $t\left(\nu,\omega\right)$ from seller ν to buyer ω, $BC\left(\nu,\omega\right)$ is the Bhattacharyya coefficient, ι_{ν} and ι_{ω} are seller and buyer fixed effects. Importantly, instead of having the usual measure for geographic distance, we consider an origin $\times$ destination geographic district fixed effect $\iota_{od}\left(\nu,\omega\right)$. Because our focus is on cultural proximity, this fixed effect helps control for features beyond the geographic distance that might arise between a pair of locations, such as cultural ties, historical ties, terrain, etc.

Columns 1 and 2 of Table 1 present the results of the intensive margin estimation, which confirm the preliminary findings from Figure 2. Column 1 shows that trade between firms

[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 $BC\left(\nu,\omega\right)$. 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 $BC\left(\omega,\nu\right)$. Regardless of the roles the firms take in this second transaction, we want their cultural proximity to remain constant, as the membership of cultural groups is fixed. This goal is achieved through the means of a symmetric proximity measure, and the Bhattacharyya coefficient complies with this symmetry requirement, as $BC\left(\nu,\omega\right)=BC\left(\omega,\nu\right)$.

Figure 2: Effect of cultural proximity on trade, intensive margin

(a) Sales (b) # Transactions



Notes: Results residualized of seller fixed effects, buyer fixed effects, and origin × destination fixed effects. Equally distanced bins formed over the horizontal axis. The size of the bubbles represents the number of transactions in each bin. The higher the Bhattacharyya coefficient, the more culturally closer the two firms are.

Table 1: Effect of cultural proximity on trade, intensive and extensive margins

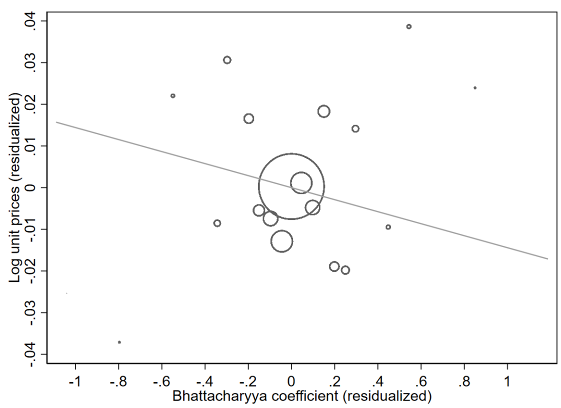
	(1)	(2)	(3)
Dep. Variable	Log Sales	Log Transactions	Trade Indicator
BC	0.129***	0.076***	0.0010***
	(0.034)	(0.028)	(0.0001)
Obs.	32,843	32,843	5,628,290
Adj. R2	0.410	0.356	0.0106
FE	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.

Notes: Columns 1 and 2 show the results of estimating Equation 1. Column 3 shows the result of estimating Equation 2. ***, ** and * indicate statistical significance at the 99, 95, and 90 percent levels 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 the two firms are. The number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia et al., 2019).

that are culturally identical will be 12.9 percent higher than trade between firms that are completely different in cultural terms. Meanwhile, Column 2 shows that the number of transactions between culturally identical firms will be approximately 7.6 percent higher than that of culturally different firms.

Fact 2: Cultural proximity increases trade at the extensive margin. 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 3: Effect of cultural proximity on prices



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

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

$$tr(\nu, \omega) = \iota_{\nu} + \iota_{\omega} + \iota_{od}(\nu, \omega) + \delta BC(\nu, \omega) + \varepsilon(\nu, \omega, t). \tag{2}$$

Column 3 of Table 1 presents the extensive margin results. We find that culturally identical firms are 0.1 percentage points more likely to trade when compared to culturally different firms. To put this coefficient into context, this means that going from a case of completely distant firms to a case evaluated at the mean of the dependent variable (0.18 percent) results in 56 percent more matches.

Fact 3: Cultural proximity lowers prices. So far, we have shown that cultural proximity affects quantities and pairings. However, given the granularity and uniqueness of our dataset, we can also study which effect cultural proximity has on unit prices.

In that sense, Figure 3 now uses buyer-seller-product-month groups and shows the residualized scatterplots between the similarity measure and 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 transaction level data and estimate

$$\ln p_g(\nu,\omega,t) = \iota_{\nu g} + \iota_{gt} + \iota_{\omega} + \iota_{od}(\nu,\omega) + \delta BC(\nu,\omega) + \epsilon_g(\nu,\omega), \tag{3}$$

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

We present the results in Table 2, which confirms the previous findings from Figure 3: the culturally closer, the lower the unit value of the transactions. Culturally identical firms will charge prices that are approximately 4 percent lower with respect to what they would charge to culturally different firms.

Table 2: Effect of cultural proximity on prices

	(1)	(2)	(3)
Dep. Variable	Log Prices	Log Prices	Log Prices
BC	-0.045*	-0.040*	-0.039*
	(0.023)	(0.023)	(0.022)
Obs.	235,001	236,617	230,900
Adj. R2	0.933	0.925	0.936
FE	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 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 more culturally closer the two firms are. The number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia et al., 2019).

3.1 Differentiated goods and contract enforcement

To understand the underlying forces driving these empirical patterns, in this section, we estimate empirical specifications where we leverage variation in differentiated goods and court quality across Indian districts. We confirm the role of cultural proximity on firm-to-firm trade in alleviating frictions originated by low-quality institutions.

The lack of contract enforcement in developing countries inhibits trade as sellers or buyers may not comply with the terms of the contract. For instance, the buyer could hold up the seller by withholding payment after the buyer receives the shipped goods. Differentiated or relationship-specific goods are subject to more severe hold-up problems, as there are fewer alternative buyers and sellers of such products, and buyers may withhold payment knowing that these goods are not useful outside of the relationship. In that sense, differentiated goods rely on better contract enforcement.

Contract enforcement can be either formal (e.g. courts) or informal (e.g. cultural proximity). Most of the literature has focused on the role of formal institutions in enforcing contracts (Boehm and Oberfield, 2020). We hypothesize and show evidence that cultural proximity alleviates contracting frictions for differentiated goods when formal contract enforcement is lacking (Nunn, 2007).

To bring in information about the type of product, we disaggregate our data at the transaction level. Then, we classify the goods into differentiated goods and non-differentiated goods based on the classification developed by Rauch (1999).[11] We estimate:

$$\ln n_g(\nu,\omega,t) = \iota_{\nu g} + \iota_{gt} + \iota_{\omega} + \iota_{od}(\nu,\omega) + \delta BC(\nu,\omega) + \xi\left(BC(\nu,\omega)\times \mathbb{I}_g^{diff}\right) + \epsilon_g(\nu,\omega), \tag{4}$$

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] Columns 2 and 3 of Table 3 present the results. Our findings suggest that the baseline results of cultural proximity increasing trade are mostly driven by differentiated goods.

To understand the channel for why trade in differentiated goods depends on cultural proximity, we turn to the analysis of contract enforcement. We posit that, when facing poor contract enforcement and low-quality institutions, firms trading differentiated goods must rely on alternative mechanisms that substitute the formal ones. Here, cultural proximity

[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.

Table 3: Effect of cultural proximity on trade by types of good, intensive margin

	(1)	(2)	(3)
Dep. Variable	Log Sales	Log Sales	Log Sales
BC	0.069**	-0.019	0.013
	(0.027)	(0.048)	(0.038)
$BC \times \mathbb{I}_g^{diff,con}$		0.139**	
		(0.059)	
$BC \times \mathbb{I}_g^{diff,lib}$			0.095**
			(0.047)
Obs.	177,584	177,584	177,584
Adj. R2	0.853	0.853	0.853
FE	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 4. ***, ** and * indicate statistical significance at the 99, 95, and 90 percent levels respectively. Good g is defined according to the 6-digit HS classification. Sales were 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 are 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. The number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia et al., 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).

arises as a substitute for the trust and enforcement a well-functioning contract would have provided (Munshi, 2014, 2019).

Court quality. To test this channel in our firm-to-firm setting, we use data from Ash et al. (2021), and calculate the average number of months between the filing of a case and its first hearing in each district court between 2010 and 2018. Intuitively, the longer the delays, the worse the contract enforcement. We estimate the following specification:

$$\begin{aligned}\ln n_g(\nu,\omega,t) = \quad & \iota_{\nu g} + \iota_{gt} + \iota_{\omega} + \delta BC(\nu,\omega) + \xi_1 \left(BC(\nu,\omega) \times \mathbb{I}^{court}(\nu,\omega)\right) \\ & + \xi_2 \left(BC(\nu,\omega) \times \mathbb{I}^{court}(\nu,\omega) \times \mathbb{I}_g^{diff}\right) + \eta \ln dist(\nu,\omega) + \epsilon_g(\nu,\omega),\end{aligned} \tag{5}$$

where $\mathbb{I}^{court}(\nu,\omega)$ is an indicator that equals 1 whenever the sum of the delays in the origin-district court, and the destination-district court is above the 75th percentile. That is, the variable indicates if a given transaction involves districts with poor contract enforcement.[13]

[13]In this specification we cannot consider the origin × destination fixed effects, as they would absorb the

Table 4: Effect of cultural proximity on trade interacted by court quality, intensive margin

	(1)	(2)	(3)
Dep. Variable	Log Sales	Log Sales	Log Sales
BC	0.051*	0.053**	0.053*
	(0.027)	(0.027)	(0.027)
$BC \times \mathbb{I}^{court}$	0.160*	0.033	0.002
	(0.094)	(0.117)	(0.114)
$BC \times \mathbb{I}^{court} \times \mathbb{I}_g^{diff,con}$		0.229*	
		(0.130)	
$BC \times \mathbb{I}^{court} \times \mathbb{I}_g^{diff,lib}$			0.273**
			(0.124)
Obs.	166,448	166,448	166,448
Adj. R2	0.851	0.851	0.851
FE	Seller×HS, buyer, month×HS	Seller×HS, buyer, month×HS	Seller×HS, buyer, month×HS

Notes: This table shows the results of estimating Equation 5. ***, ** and * indicate statistical significance at the 99, 95, and 90 percent levels respectively. Good g is defined according to 6-digit HS classification. Sales were trimmed by 4-digit HS code at 5 and 95 percent. Standard errors two-way clustered at the seller and 4-digit HS level. Standard errors in parentheses. The higher the Bhattacharyya coefficient, the more culturally closer the two firms are. The number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia et al., 2019). $\mathbb{I}^{court}(\nu, \omega)$ indicates if the sum of the delays in the origin-district court and the destination-district court is above the 75th percentile. $\mathbb{I}_g^{diff,con}$ 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).

Column 1 of Table 4 shows that, while cultural proximity is important for firm-to-firm trade overall, it is particularly relevant for those pairs of districts with low court quality. We interpret this as evidence that firms rely on cultural proximity as a source of trust in places where institutions do not work well.

Importantly, Columns 2 and 3 of Table 4 show that such an effect is primarily explained by differentiated goods. In places where contract enforcement is poor, cultural proximity allows the transaction of the type of goods that rely the most on strong contracts and institutions. This provides evidence for our proposed explanation that cultural proximity acts as an informal substitute for the trust that a well-functioning contract would provide (Munshi, 2019; Nunn, 2007).

Differentiated goods are branded and specific to certain producing firms. In a country with market imperfections, firms can easily renege on their commitments. This is particularly

origin × destination court quality variation we are interested in. Thus, we consider the Euclidean origin × destination distance $dist(\nu, \omega)$ as a control variable.

exacerbated in regions with low court quality and low contract enforcement. Unlike homogeneous goods, firms in differentiated good markets are not easily replaceable. As a result, firms buying or selling differentiated goods will only trade with firms they know and trust, and perhaps are culturally closer.[14] As we show, this is particularly exacerbated in areas where the court system is delayed and backed up.

3.2 Additional specifications and understanding mechanisms

In Appendix C, we examine alternative specifications and heterogeneity in responses that shed light on various channels and mechanisms. Here we summarize the findings.

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 A3 and A4 show our baseline findings are robust to this alternative cultural proximity measure.

Caste hierarchies. As a way for testing for preference-based discrimination across the social hierarchy, we analyze if there are asymmetric effects in transactions where one firm is placed higher than the other based on the Varna-based social order. Table A5 shows no additional effect of cultural proximity when firms are placed differently in the social hierarchy. While vertical hierarchies and preference-based discrimination are prevalent in Indian society, we do not detect that vertical hierarchies are what governs trade between firms.

Goods specialization. Cultural groups in India may be involved in particular occupations, and so as a result, in the production of specific goods (Munshi, 2019). Therefore, we analyze if the reason behind the cultural proximity results are cultural groups specializing in the production of certain goods and, given this, forming special bonds with their specific set of buyers. In Table A6 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.

Number of varieties sold and bought. We analyze whether firms' dependence on cultural proximity for trade prevails for firms that sell and buy more varieties of goods. We hypothesize 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 clients. These firms, to minimize their load of contracting frictions, will rely more on trading with those that are culturally

[14] Rauch (1999) argues 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.

closer. We count how many varieties of inputs a firm buys or how many varieties of goods a firm sells. In Table A7, we indeed find the more varieties a firm sells or buys, the more the intensity of trade is affected by cultural proximity.

Age of firms. We use the age of firms as a way to test for taste-based discrimination. The idea is that, if there are taste-based preferences, then firms who sell to culturally close firms at lower prices are willing to forego profits because of their preferences (Becker, 1957). As a result, it is more likely these firms will go bankrupt and exit the market. We use the age of the firms as a proxy for which firms have survived longer. Table A8 shows the results of the intensive margin regressions after controlling for age. If there was taste-based discrimination, then the interaction between the measure of cultural proximity and age should have a negative coefficient: older firms should rely less on cultural proximity (i.e. they have been able to reach old age because of this). Nonetheless, we detect only weak evidence for taste-based discrimination. As such, we cannot establish that taste-based discrimination is the channel that drives our cultural proximity results.

Industry pair linkages. In the production matrix of an economy, some sectors are more likely to trade with others because of the nature of their activities. The same cultural group may happen to participate in the same industry. In Table A9, we present the results for the intensive margin after adding an industry of seller $\times$ industry of buyer fixed effect. We find that the result of there being more trade between culturally closer firms prevails.

Cancellations. In line with our discussion on contract enforcement, we recognize that cultural proximity can affect trade by making contract reneging less likely. A proxy for this in our dataset is the number of canceled transactions between firms. In Table A10, we find that there is a lower probability of a transaction being canceled when the firms involved are close in cultural terms. While not conclusive, as we do not measure reneged contracts directly, this result is consistent with our primary narrative of contract enforcement issues. This again suggests that cultural proximity matters by lowering the rate of contract reneging.

Correction for selection bias. Helpman et al. (2008) shows that the standard gravity equation estimations are biased as they do not account for selection issues. We follow their suggested correction in Tables A11 and A12. As the correction mentions, we need an excluded instrument that affects only the extensive margin (i.e. the matching cost) and not the intensive margin (i.e. the trade cost). We consider the participation of both seller and buyer on IndiaMART, India's largest online B2B platform, under the idea that online platforms reduce their matching costs. The results show that the coefficients are downward biased if we do not account for the selection issues. Therefore, our main results in the paper provide a lower bound on the effect that cultural proximity has on the intensive margin of trade.

3.3 Discussion of stylized facts

The stylized facts show that higher cultural proximity between a pair of firms fosters trade at both intensive and extensive margins. We discuss the possible mechanisms that may give rise to these findings.

Contracting frictions. In Section 3.1, we argue that contracting frictions could be the reason that drives the cultural proximity results. India is a country that suffers from a 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, 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 partners. Both on the extensive and intensive margin, this lowers trade as firms must pay a matching and trade cost. This increases prices as the costs are passed down by the sellers to the aforementioned prices.

In such a situation, cultural proximity can work as a proxy for information and trust: culturally closer firms may know and trust each other, and informally enforce contracts with social and reputational pressures. The higher the cultural proximity, the lower the contracting frictions. As a result, there would be more trade and lower prices, consistent with our findings. In Section 4, we present a simple theoretical framework in which cultural proximity affects contracting frictions and affects trade and prices. Our model suggests that if contracting frictions drive initial trade barriers, cultural proximity reduces such frictions.

Preference-based mechanisms and discrimination. We argue our results are unlikely to emerge from buyers having an inherent preference for buying from sellers culturally closer to them. This preference would be 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 inconsistent with our previous findings. Indeed, the ability to measure prices in our dataset allows us to distinguish between various channels.

The stylized facts can arise from having sellers that show a preference for selling to culturally closer buyers. This would imply the introduction of a supply shifter that is active for those buyers who are culturally closer to the seller. However, we do not find conclusive evidence of this channel in Section C.5.

Discrimination from high-caste cultural groups against low-caste cultural groups may again reduce trade. Yet, in Section C.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.

4 Aggregate effects of cultural proximity

Our goal in this section is to propose a general equilibrium model to quantify the aggregate effects of cultural proximity. We start with a static production network model, and allow firm owners to differ in their cultural endowments, which we then use to construct measures of cultural proximity within a production network. We then connect the model with our detailed firm-to-firm dataset and micro-level stylized facts. Finally, we perform counterfactual exercises that quantify the aggregate effects of cultural proximity and the quality of courts.

4.1 Setup

Preferences and technology. A representative household demands goods from firms and inelastically supplies labor to firms. A continuum of firms operate under monopolistic competition and produce differentiated goods. The economy is a roundabout production economy, so firms can purchase intermediate inputs from all firms in the economy. Each firm produces its differentiated good with hired labor from the representative household and purchased intermediate inputs.

Firms operate in three steps. First, sellers endogenously choose buyers. Second, upon matching, sellers charge a premium to buyers arising from contracting frictions. Third, sellers and buyers trade.

Cultural proximity in production networks. Guided by our stylized facts, cultural proximity shapes trade in two ways. First, cultural proximity between sellers and buyers reduces the cost of matching, and so affects the extensive margin of who trades with whom. To model this, we posit that culture encodes information that is useful for sellers to decide which buyer to sell to. Second, given the risk that buyers may renege on payment under the low-quality institutional setting, sellers determine the premium charged to buyers, which affects the intensive margin of trade.

4.2 Preferences

A representative household demands goods from firms with CES $\sigma > 1$, the elasticity of substitution between goods, and inelastically supplies labor to firms. The household maximizes utility

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

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

$$x(\omega) = P(\omega)^{1-\sigma} P^{\sigma-1} Y, \tag{6}$$

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

4.3 Technology

There is a continuum of firms that operate under monopolistic competition and produce differentiated goods indexed by ω. Since each firm produces a unique good, ω denotes a firm or a good. We consider a roundabout production economy, so firms can purchase intermediate inputs from all firms in the economy. Each firm produces its differentiated good with hired labor from the representative household and intermediate inputs.

Firms operate in three steps. First, sellers endogenously choose buyers (i.e. matching). In particular, a seller ν sells to buyer ω whenever the profits of doing so are larger than the fixed costs of matching $\epsilon F(\nu, \omega)$. Guided by our stylized facts, we argue that cultural proximity $BC(\nu, \omega)$ between seller ν and buyer ω determines the pairwise matching costs $F(\nu, \omega)$. Indeed, previous work describes how culture encodes information that is useful for sellers to decide which buyer to sell to (Allen et al., 2019; Balmaceda and Escobar, 2017).[15]

Second, upon matching, due to the risk of buyers withholding payment, sellers charge a premium to buyers (i.e. terms of trade). In particular, the price that seller ν charges to buyer ω includes a premium $d(\nu, \omega) \geq 1$, which is a function of different factors that influence the hold-up process. Guided by our stylized facts, we posit that cultural proximity $BC(\nu, \omega)$ between seller ν and buyer ω determines $d(\nu, \omega)$. This is similar to earlier work

[15] There could be other microfoundations motivating how culture influences matching costs (e.g. risk-sharing as in Ambrus et al. 2014; Bloch et al. 2008. These would generate similar frictions at the extensive and intensive margin.

on how culture can be an informal institutional channel to solve hold-up problems.[16] Third, sellers and buyers trade.

Step 1: Matching. In the first step, we endogenize the formation of the production network by laying out the maximization problem of firms, and how cultural proximity influences it. A seller ν matches with buyer ω whenever the seller's profits of doing so $\pi(\nu, \omega)$ are larger than the fixed costs of matching $\epsilon F(\nu, \omega)$, where $F(\nu, \omega)$ are pairwise fixed costs of matching, and ϵ are i.i.d. log normal errors with mean $\mu_{\ln(\epsilon)}$ and standard deviation $\sigma_{\ln(\epsilon)}$. It can be shown that profits are proportional to the value of intermediate sales $n(\nu, \omega)$ from seller ν to buyer ω, such that

$$\pi(\nu, \omega) = \frac{n(\nu, \omega)}{\sigma}. \tag{7}$$

Sellers are looking for buyers, and vice-versa. Firms can meet costlessly, so all firms can meet each other before deciding who to match with. The purpose of a meeting is to infer the fixed costs the seller would incur if they were to trade with each buyer. Here we rationalize how sellers infer the costs of matching arising from cultural proximity. We assume that each firm owner has a name and that sellers and buyers exchange names upon meeting. Then, each name has a probability mapping to different cultural groups. We assume that this mapping is public information. So, for example, we assume that all firm owners know that the surname *Shah* is associated with either of the three groups: Jain, Muslim (Faqir), or Hindu (Vaishnav baniya), with probabilities reflecting the empirical distribution.

For each cultural group j, we consider the matching function

$$M^j(\nu, \omega) = (\rho_\nu(j))^\varphi (\rho_\omega(j))^{1-\varphi},$$

where $\rho_\nu(j)$ is the probability that seller ν belongs to cultural group j, $\rho_\omega(j)$ is the probability that buyer ω belongs to cultural group j, φ is the weight of ν to determine their proximity. For simplicity, we assume $\varphi = \frac{1}{2}$. The expected proximity between ν and ω is

$$\overline{M}(\nu, \omega) = \frac{1}{\mathcal{X}} \sum_{j=1}^{\mathcal{X}} M^j(\nu, \omega) \propto BC(\nu, \omega).$$

That is, the expected proximity between seller ν and buyer ω is proportional to the Bhattacharyya coefficient we use in the empirical part of the paper to measure cultural proximity between firms. Finally, following our stylized facts, the fixed costs of matching are a function

[16] Again, this would be consistent with other microfoundations on how culture influences trade costs due to reputation (Banerjee and Duflo, 2000; Chen and Wu, 2021) or loyalty (Board, 2011).

of cultural proximity such that $F(\nu,\omega) = f\left(\overline{M}(\nu,\omega)\right)$, where $f(\cdot)$ is an exponential function for simplicity. Our modeling decision allows for the fact that a firm belonging to a cultural group encodes information that other firms use to determine the cost of matching. Then, we consider

$$F(\nu,\omega) = \kappa + \exp\left(-\gamma BC(\nu,\omega)\right), \tag{8}$$

where $\gamma > 0$ measures the sensitivity of the pairwise matching cost to the cultural proximity $BC(\nu,\omega)$, and κ is a scaling constant.

The intuition behind parameter γ is that it encases court quality, such that $\gamma = \gamma\left(\mathbb{I}^{court}\right)$, where $\mathbb{I}^{court}$ is an indicator of low court quality. In this sense, Table A1 in Appendix A suggests that parameter γ increases in magnitude when $\mathbb{I}^{court} = 1$ and firms face weak court quality. The lower the quality of courts is, the larger parameter γ is, and the more firms rely on cultural proximity to match.

We now describe how firms match. Before the formation of the network, firms are characterized by $\lambda = (z, \boldsymbol{\rho})$, where z is productivity, and $\boldsymbol{\rho}$ is the vector of probabilities of firm λ belonging to each cultural group. After firms meet each other and infer their cultural proximity, firms now only differ in their productivity z. So the share of seller-buyer pairs $\left(z, z^{'}\right)$ is

$$l\left(z, z^{'}\right) = \int \mathbb{I}\left[\ln\left(\pi\left(z, z^{'}\right)\right) - \ln\left(F\left(z, z^{'}\right)\right) - \ln\left(\epsilon\left(z, z^{'}\right)\right) > 0\right] dH\left(\epsilon\left(z, z^{'}\right)\right), \tag{9}$$

where $l\left(z, z^{'}\right)$ is called the *link function*. From Equations 8 and 9, 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.

Step 2: Terms of the contract. After matching, seller ν and buyer ω set up the terms of the contract. We first consider the possibility that the buyer may withhold payment after the seller ships the goods. The seller's concerns around this issue can be monetized in the premium that the seller charges to the buyer. This premium gets passed to the unit price $p(\nu,\omega)$ that results from profit maximization

$$p(\nu,\omega) = \mu c(\nu)\, d(\nu,\omega), \tag{10}$$

where $c(\nu)$ is the marginal cost and $\mu \equiv \frac{\sigma}{\sigma-1}$ is a markup.

We now turn to explain the premium $d(\nu,\omega)$. Time is continuous within our one-period static model. Seller ν considers the time T that buyer ω withholds payment for the first time is a random variable. T follows an exponential distribution with intensity rate $\delta(\nu, \omega)$. Then, the probability seller ν waits for more than t units of time until buyer ω withholds payment for the first time is

$$\begin{aligned} P(T > t) &= 1 - P(T \leq t), \\ &= 1 - (1 - \exp(-\delta(\nu, \omega) t)), \\ &= \exp(-\delta(\nu, \omega) t). \end{aligned}$$

Seller ν cares that buyer ω never withholds payment. Then, the probability $\overline{p}$ that seller ν waits more than a unit of time until buyer ω withholds payment is

$$\begin{aligned} \overline{p} &= P(T > 1), \\ &= \exp(-\delta(\nu, \omega)). \end{aligned}$$

Then, the expected number of times buyer ω withholds payment is $\frac{1}{\overline{p}}$. We posit that the premium $d(\nu, \omega)$ is proportional to the expected number of times buyer ω withholds payment to seller ν. The intuition is that seller ν will charge a higher premium if there is a higher hold-up risk from buyer ω. For simplicity,

$$\begin{aligned} d(\nu, \omega) &= \frac{1}{\overline{p}} \geq 1, \\ &= \exp(\delta(\nu, \omega)). \end{aligned}$$

Finally, we allow $\delta(\nu, \omega)$ to include a set of covariates that affects the terms of trade. For example, inputs can spoil or get lost more easily with larger distances, which increases the probability of buyer ω not paying seller ν since the goods did not arrive as agreed.

More importantly, guided by our stylized facts, we allow for the fact that, upon matching, cultural proximity $BC(\nu, \omega)$ also influences $\delta(\nu, \omega)$. In particular, we consider

$$d(\nu, \omega) = \exp(\beta BC(\nu, \omega) + \epsilon(\nu, \omega)), \tag{11}$$

where $\beta < 0$ is a trade cost semi-elasticity, and $\epsilon(\nu, \omega)$ are unobservables.

From Equation 10, we have that the higher the cultural proximity, the lower the prices, which relates to Stylized Fact 3. Likewise, from Equation 15, we see that the higher the cultural proximity, the higher the intermediate sales, which relates to Stylized Fact 1.

The economic intuition of parameter β is that it incorporates information on court quality, such that $\beta = \beta(\mathbb{I}^{court})$. Based on the result of Table 4, we know model parameter β increases in magnitude when $\mathbb{I}^{court} = 1$ and firms face bad courts. Therefore, parameter β relates to how firms respond to court quality: the lower the quality of the courts, the larger the magnitude of parameter β, and the more important cultural proximity is for trade.

Step 3: Trade. Finally, we model trade. Each firm has a technology

$$y(\omega) = \kappa_\alpha z(\omega) \, l(\omega)^\alpha \, m(\omega)^{1-\alpha}, \tag{12}$$

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

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

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

$$c(\omega) = \frac{P(\omega)^{1-\alpha}}{z(\omega)}, \tag{13}$$

where $P(\omega) \equiv \left(\int_{\nu \in \Omega(\omega)} p(\nu, \omega)^{1-\sigma} \, d\nu \right)^{\frac{1}{1-\sigma}}$ 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 the unit price is $p(\nu, \omega) = \mu c(\nu) \, d(\nu, \omega)$, as stated before.

We now derive the demand for intermediates, so

$$n(\nu, \omega) = p(\nu, \omega)^{1-\sigma} \, P(\omega)^{\sigma-1} \, N(\omega), \tag{14}$$

where $N(\omega) = \int_{\nu \in \Omega(\omega)} n(\nu, \omega) \, d\nu$ is the total intermediate purchases by buyer ω and $n(\nu, \omega) \equiv p(\nu, \omega) \, m(\nu, \omega)$ is the value of purchases from seller ν to buyer ω. From Equation 14 we can obtain the gravity equation as

$$\log(n(\nu, \omega)) = \iota_\nu + \iota_\omega + (1 - \sigma) \log(d(\nu, \omega)), \tag{15}$$

where ι_ν and ι_ω are seller and buyer fixed effects. Here, the premium $d(\nu, \omega)$ enters the gravity equation as a trade cost. This gravity equation relates directly to Equation 1 that we estimate.

4.4 Equilibrium given production network

Conditional on the network structure, firms only differ in productivity z. Based on the price index of all of the goods acquired by firm $z^{'}$, we get

$$P\left(z^{'}\right)^{1-\sigma}=\mu^{1-\sigma}\int P\left(z\right)^{(1-\alpha)(1-\sigma)}z^{\sigma-1}d\left(z,z^{'}\right)^{1-\sigma}l\left(z,z^{'}\right)dG\left(z\right), \tag{16}$$

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 households plus intermediates, so

$$S\left(z\right)=\frac{\left[\mu^{1-\sigma}P\left(z\right)^{(1-\alpha)(1-\sigma)}z^{\sigma-1}\right]\times}{\left[\frac{Y}{P^{1-\sigma}}D\left(z\right)^{1-\sigma}+\left(\frac{1-\alpha}{\mu}\right)\left(\int\left[d\left(z,z'\right)^{1-\sigma}P\left(z'\right)^{\sigma-1}S\left(z'\right)\right]l\left(z,z'\right)dG\left(z'\right)\right)\right]}, \tag{17}$$

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.

5 Estimation and calibration

We explain how we estimate the key parameters of the model on cultural endowments, trade costs, matching costs, and the remaining parameters of the model.

Cultural endowments $\boldsymbol{\rho}$. For the cultural endowments, we assume each firm ν has a probability vector $\boldsymbol{\rho_\nu}=\left[\rho_\nu\left(1\right),\ldots,\rho_\nu\left(452\right)\right]$ 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 $\rho_\nu\left(1\right),\ldots,\rho_\nu\left(452\right)\sim\mathcal{D}\left(\alpha_1,\ldots,\alpha_{452}\right)$, where $\alpha_1,\ldots,\alpha_{452}>0$ are concentration parameters.[17] The probability density for the Dirichlet distribution is

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

such that $\rho_\nu\left(x\right)\in\left[0,1\right],\quad\sum_{x=1}^{452}\rho_\nu\left(x\right)=1$, where $\Gamma\left(.\right)$ 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

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

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.[18] 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

$$\ln pr\left(\boldsymbol{\varrho}|\boldsymbol{\alpha}\right) = \mathcal{N} \ln \Gamma\left(\sum_{x=1}^{452} \alpha_x\right) - \mathcal{N} \sum_{x=1}^{452} \ln \Gamma\left(\alpha_x\right) + \mathcal{N} \sum_{x=1}^{452} \left(\alpha_x - 1\right)\left(\frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \ln \rho_\nu\left(x\right)\right). \quad (18)$$

Trade cost d. From Equation 11 we need an estimate for β. Our setup produces an empirical counterpart we estimated in the reduced-form section. That is, we obtain estimates for this parameter by linking the theoretical gravity Equation 15 to the empirical gravity equation results (Column 1 from Table 1). Given $\sigma = 3.94$, which we discuss below, we obtain $\beta = -0.04$.[19]

Matching cost F. From Equation 8, we need an estimate for γ. We do this in two steps. First, using the extensive margin sample we estimate the following specification:

$$\ln\left[n\left(z, z^{'}\right)\right] = \iota_z + \iota_{z^{'}} + \iota_{od}\left(z, z^{'}\right) + \varepsilon\left(z, z^{'}\right), \quad (19)$$

where we apply the inverse hyperbolic sine transformation to the dependent variable, to not lose the cases in which there is zero trade. The seller and buyer fixed effects ι_z and $\iota_{z^{'}}$ absorb productivity, size and other firm characteristics. Also, we capture distance and other variables that vary by geography with the origin $\times$ destination fixed effect $\iota_{od}\left(z, z^{'}\right)$.

Then, we recover

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

where the hats are predicted fixed effects and $\ln\widehat{\left[n\left(z, z^{'}\right)\right]}$ are the predicted sales.

Next, we combine and rearrange Equations 8 and 9, such that

$$l\left(z, z^{'}\right) = \int 1\left[\ln\left(\epsilon\left(z, z^{'}\right)\right) < \ln\widehat{\left[n\left(z, z^{'}\right)\right]} - \ln\left(\sigma\right) - \gamma BC\left(z, z^{'}\right)\right] dH\left(\epsilon\left(z, z^{'}\right)\right), \quad (20)$$

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 $\ln\widehat{\left[n\left(z, z^{'}\right)\right]}$.[20] Since $\epsilon\left(z, z^{'}\right)$ is log normally distributed, we estimate this last

[18] We present the estimated parameters in Figure A1 in Appendix A.

[19] Even though the wedge also appears in the price Equation 10 of the model, we do not estimate this equation to identify β. This is because the price equation is not an equilibrium statement, while the gravity equation is.

[20] We ignore the scaling constant κ that appears in Equation 8.

equation with a probit regression. We find that $\gamma = 0.17$, and statistically different from 0.[21]

Calibrated parameters and SMM. We calibrate the labor cost share $\alpha = 0.34$, the value reported for India for 2019 from the Penn World Tables (Feenstra et al., 2015). 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 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 target the average log outdegree (the number of buyers a seller has) and the log indegree (the number of sellers a buyer has), where we normalize both measures by the total number of firms. Since these two moments are related to the 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, to identify the standard deviation of the link function noise distribution $\sigma_{\ln(\epsilon)}$ we target the variance of the log outdegree. Lastly, to calibrate the standard deviation of the log productivity $\sigma_{\ln(z)}$, we must choose a moment that is related to the variance of the intensive margin. Thus, we target the variance of the log intermediate sales, where we normalize the sales with respect to the total.

The first untargeted moment we consider is the variance of the log indegree. The second untargeted moment we examine is the variance of the log prices of intermediate goods. We further discuss these moments in Appendix B.

Goodness of fit. We estimate the parameters $\sigma_{\ln(z)} = 0.85$, $\mu_{\ln(\epsilon)} = 65.14$, $\sigma_{\ln(\epsilon)} = 10.98$ and $\kappa = 15.36$. Table A2 in Appendix B shows how the model-based moments fare against their empirical counterparts. When it comes to the targeted moments, the model can very

[21] We present the results of the estimation in Table A1 in Appendix A. We also present a specification where we account for court quality.

closely replicate the empirical ones. For the untargeted moments, the model gets reasonably close to the data.

6 Counterfactuals

We now turn to the aggregate effects of cultural proximity. First, we quantify the extent and different aspects by which cultural proximity shapes the economy. To do this, we compare an economy where firm owners are culturally dissimilar with one where firm owners are culturally homogeneous.

Second, we quantify the aggregate extent of the distortions that cultural proximity introduces into the real-world economy with respect to a frictionless one. To do this, we study a case where we start with an economy where firms are culturally the closest and move to the baseline case, where cultural proximity is disciplined by the data.

Third, we relate our results directly to policy, and analyze the effects of improving court quality. Based on our empirical findings, the trade cost and the matching cost are more sensitive to cultural proximity when firms face low-quality courts. Therefore, we introduce a policy that improves the quality of courts in our model by reducing the magnitude of parameters β and γ.

To quantify the aggregate effects of cultural proximity under each scenario, we measure various model-based aggregate 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 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, for the extensive margin, we consider the average of all the elements of the link function $l\left(z, z'\right)$. This measures the average probability of matching of any given firm in the economy. Lastly, for the prices, we compare changes in the aggregate price index P.

CF1: From culturally distant to culturally close

In this first counterfactual exercise, we quantify the aggregate effect of going from an economy where firms are culturally the farthest to an economy where firms are culturally the closest. We designed this counterfactual exercise to show the potential aggregate effects of

Table 5: Effect of cultural proximity on aggregate outcomes (counterfactual scenarios)

	CF1: From $BC\left(z,z'\right)=0$ to $BC\left(z,z'\right)=1$	CF2: From $BC\left(z,z'\right)=1$ to baseline	CF3: Improving courts in baseline
Welfare	10.85%	-6.20%	7.71%
Ave. productivity	-0.22%	0.06%	-0.10%
Total sales	11.04%	-7.32%	9.31%
Ave. prob of matching	0.09p.p.	-0.04p.p.	0.05p.p.
Agg. price index	-9.79%	6.61%	-7.15%

Notes: CF1 is a scenario where we go from a world where all firms are completely distant ($BC\left(z,z'\right)=0$ for all $z\neq z'$) to a frictionless world ($BC\left(z,z'\right)=1$ for all z,z'). CF2 is a scenario where we go from a frictionless world to the baseline world (where we draw realizations of cultural endowments according to Section 5). CF3 is a scenario where courts improve under the baseline data-based draws of $BC\left(z,z'\right)$.

Table 6: Change in sales by productivity quartiles

Case	CF1: From $BC\left(z,z'\right)=0$ to $BC\left(z,z'\right)=1$	CF2: From $BC\left(z,z'\right)=0$ to baseline	CF3: Improving courts in baseline
1st quartile (most productive)	10.94%	-7.27%	9.24%
2nd quartile	11.55%	-7.55%	9.65%
3rd quartile	11.76%	-7.78%	9.93%
4th quartile (least productive)	12.10%	-8.05%	10.24%

Notes: CF1 is a scenario where we go from a world where all firms are completely distant ($BC\left(z,z'\right)=0$ for all $z\neq z'$) to a frictionless world ($BC\left(z,z'\right)=1$ for all z,z'). CF2 is a scenario where we go from a frictionless world to the baseline world (where we draw realizations of cultural endowments according to Section 5). CF3 is a scenario where courts improve under the baseline data-based draws of $BC\left(z,z'\right)$.

cultural proximity to their full extent. Specifically, this exercise involves changing the cultural proximity under the standard calibration. We start from $BC\left(z,z'\right)=0$ for all $z\neq z'$ and go to a case where $BC\left(z,z'\right)=1$ for all z,z'.

Table 5 shows how the model statistics change in each counterfactual. In case CF1, we have that firms become the closest in cultural terms. Aligned with our empirical facts, with lower trade costs, total sales increase by 11.04 percent. With the lower matching costs, the average probability of matching increases by 0.09 percentage points. Also, because there are lower trade and matching costs, aggregate prices fall by 9.79 percent. With this, welfare increases by 10.85 percent.

Yet, culture-based trade comes with certain nuances, as it is not necessary that the most productive firms are more likely to trade. We find that average productivity falls by 0.22 percent. Average productivity depends on substantial compositional changes, as less productive firms are selling more or less with respect to the starting point. We show in Table 6 that when trade and matching costs decrease, the less productive firms sell more, which increases their weight in the aggregate and lowers average productivity.

The results of CF1 show that, overall, there are gains in terms of welfare, sales, and matching of moving to a world where cultural distance is non-existent (i.e. everyone becomes culturally the closest). However, it also generates drawbacks in terms of productivity: less-productive firms are trading, whereupon they would not otherwise.

CF2: How does cultural proximity distort the real economy?

In this second counterfactual exercise, we quantify the aggregate distortions introduced by cultural proximity as in the data by comparing this scenario to a frictionless one (i.e. an economy where firm owners are culturally the closest).

This exercise involves going from a frictionless case where $BC\left(z, z^{'}\right) = 1$ for all $z, z^{'}$ to a case where we discipline $BC\left(z, z^{'}\right)$ in a data-based way according to Section 5. In both cases, we keep the standard calibration constant.

In CF2, because firms become more culturally distant, trade and matching costs increase. Table 5 shows this translates into 7.32 percent fewer sales and an average probability of matching that is 0.04 percentage points lower. Because of higher costs, aggregate prices go up by 6.61 percent.

Welfare decreases by 6.20 percent. At the same time, average productivity grows by 0.06 percent. As before, this last result depends on compositional changes. Table 6 presents that in CF2 all firms face a decrease in sales. However, the least productive firms are hit the most. Therefore, overall productivity goes up.

Overall, the results of CF2 show that the real world suffers losses because of the distortions introduced by cultural proximity with respect to a frictionless world. We conclude that the net aggregate effect on trade and welfare, of having culturally distant firms is negative.

CF3: Improving court quality

We discussed how, empirically, court quality plays a relevant role in the determination of firm-to-firm trade. In this third counterfactual, we set out to quantify the aggregate effect of a policy that improves court quality, while keeping cultural proximity constant.

We fix cultural proximity $BC\left(z, z'\right)$ to be disciplined by the data according in Section 5, and reduce the magnitudes of parameters β and γ so as to resemble a policy that improves court quality. In contrast to the previous exercises, now we follow Table 4 for the calibration of parameter β. In this exercise, we set that β goes from -0.07 to -0.02 as courts improve. Also, based on the result of Table A1, we assume γ changes in this exercise from 0.26 to 0.14 when courts improve. Thus, we introduce better courts into the model as a measure that reduces the sensitivity of trade costs and matching costs to cultural proximity.

The lower sensitivity of trade costs and matching costs, while keeping cultural proximity unchanged, has a net effect of lowering these costs. This way, Table 5 shows total sales increase by 9.31 percent, and the average probability of matching goes up by 0.05 percentage points. Given these decreasing costs, aggregate prices fall by 7.15 percent, and welfare increases by 7.71 percent.

Average productivity decreases by 0.10 percent. Table 6 shows that this result is due to the least productive firms selling more after the change. As trade and matching costs fall because of the introduction of the policy, the weight of low-productivity firms grows with respect to the more productive ones, and average productivity falls.

In a setting of low-quality courts, firms will try to solve the institutional frictions by basing their trade decisions on cultural proximity. CF3 shows that, quantitatively, improving courts reduces the sensitivity of trade costs and matching costs to cultural proximity. In particular, given a constant cultural proximity between firms, better courts result in lower trade costs and matching costs, such that there are aggregate welfare gains.

7 Conclusions

Emerging economies feature low-quality institutions that generate trade frictions at the firm-to-firm level. In these economies, firms resort to informal institutions to overcome trade frictions. While at first sight, informal institutions may ease trade at the firm-to-firm level, their impact on the aggregate economy is less clear, as trade with low-productivity firms may increase.

So far, quantifying these aggregate effects has been challenging because of the lack of data on how firms trade and relate culturally, along with the prices of each transaction. In this paper, we introduce a new firm-to-firm transactions database from a large Indian state, with data on cultural proximity derived from India's caste and religious system. India is a representative example of an emerging economy that features low-quality institutions and where cultural proximity plays an important role in the economy (Munshi, 2019).

With the help of this new dataset, we report three new stylized facts that relate to cultural proximity and trade at the firm-to-firm level. First, culturally closer firms report higher sales between them, on the intensive margin. Second, firms that are culturally closer are more likely to ever trade with each other, on the extensive margin. Third, firms that are culturally closer charge lower unit prices in their transactions. This last empirical fact is particularly relevant to our analysis, as it confirms that the channel by which cultural proximity affects trade is through lowering costs.

We link these stylized facts to institutional quality. The effect we find of cultural proximity on trade is stronger for differentiated goods, which often rely on either formal or informal contract enforcement. We also find that the importance of cultural proximity is elevated in regions with low court quality (a proxy for contract enforcement and institutional quality). Indeed, the importance of court quality is only seen in the trade of differentiated products, rather than homogeneous goods. We understand these results as evidence that cultural proximity is an informal mechanism that substitutes formal contract enforcement.

While striking, the empirical facts are silent about the aggregate effects of cultural proximity. In order to quantify the aggregate effects of cultural proximity, we build a quantitative general equilibrium model of firm-to-firm trade and cultural proximity grounded on the empirical findings. We introduce our measure of cultural proximity as a wedge that affects trade and matching costs, and estimate the key parameters of the model. Importantly, we allow for institutional quality to affect the parameters that govern these wedges.

Our counterfactual exercises show that cultural proximity can have an important impact in aggregate. There are gains in terms of welfare, sales, and matching of moving to a world where cultural distance is non-existent (i.e. everyone becomes culturally the closest). However, it also generates drawbacks in terms of productivity: there are unproductive firms trading.

In another exercise, we show the effects of improving courts within our model, our proxy for institutional quality. We find that, quantitatively, improving courts reduces the sensitivity of trade costs and matching costs to cultural proximity. In particular, better courts result in lower trade costs and matching costs, such that there are aggregate welfare gains.

While there is previous evidence that cultural ties matter for trade at the cross-country level, our paper goes beyond relying on borders, and sheds new light on the micro-level mechanism for which firms may rely on cultural proximity: the lack of good institutions. When low-quality institutions prove to be a challenge for markets and trade, agents will turn to informal substitutes to ease the imposed frictions. Importantly, this micro-level evidence using granular data allows us to ground our quantitative model, with which we show that cultural proximity is not just a micro-level occurrence: it has aggregate implications in low-quality institutional settings.

Our work speaks to policy. Improving institutions can make trade less reliant on cultural proximity or, more broadly, on informal institutions. Moreover, improving contract enforcement has the potential to improve welfare by facilitating trade within firm networks.

References

Afridi, F., Dhillon, A., Li, S. X., and Sharma, S. (2020). Using social connections and financial incentives to solve coordination failure: A quasi-field experiment in india's manufacturing sector. *Journal of Development Economics*, 144:102445.

Allen, T., Bilir, K., Chen, Z., and Tonetti, C. (2019). Knowledge diffusion through networks. *Working Paper*.

Ambrus, A., Mobius, M., and Szeidl, A. (2014). Consumption risk-sharing in social networks. *American Economic Review*, 104(1):149–82.

Antras, P., Fort, T. C., and Tintelnot, F. (2017). The margins of global sourcing: Theory and evidence from US firms. *American Economic Review*, 107(9):2514–64.

Ash, E., Asher, S., Bhowmick, A., Bhupatiraju, S., Chen, D., Devi, T., Goessmann, C., Novosad, P., and Siddiqi, B. (2021). In-group bias in the Indian judiciary: Evidence from 5 million criminal cases. *Working Paper*.

Balmaceda, F. and Escobar, J. F. (2017). Trust in cohesive communities. *Journal of Economic Theory*, 170:289–318.

Bandyopadhyay, S., Coughlin, C. C., and Wall, H. J. (2008). Ethnic networks and US exports. *Review of International Economics*, 16(1):199–213.

Banerjee, A. V. and Duflo, E. (2000). Reputation effects and the limits of contracting: A study of the Indian software industry. *Quarterly Journal of Economics*, 115(3):989–1017.

Bazzi, S., Hilmy, M., and Marx, B. (2020). Religion, education, and the state. *Working Paper*.

Becker, G. S. (1957). *The economics of discrimination*. University of Chicago Press.

Bernard, A. B., Dhyne, E., Magerman, G., Manova, K., and Moxnes, A. (2022). The origins of firm heterogeneity: A production network approach. *Journal of Political Economy*, 130(7):1765–1804.

Bhagavatula, S., Bhalla, M., Goel, M., and Vissa, B. (2018). Cultural diversity on corporate boards and firm outcomes. *Working Paper*.

Bhattacharyya, A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society*.

Bloch, F., Genicot, G., and Ray, D. (2008). Informal insurance in social networks. *Journal of Economic Theory*, 143(1):36–58.

Board, S. (2011). Relational contracts and the value of loyalty. *American Economic Review*, 101(7):3349–3367.

Boehm, J. (2015). The impact of contract enforcement costs onoutsourcing and aggregate productivity. *Working Paper*.

Boehm, J. and Oberfield, E. (2020). Misallocation in the market for inputs: Enforcement and the organization of production. *Quarterly Journal of Economics*, 135(4):2007–2058.

Boken, J., Gadenne, L., Nandi, T., and Santamaria, M. (2022). Community networks and trade. *Working Paper*.

Chen, M. X. and Wu, M. (2021). The value of reputation in trade: Evidence from Alibaba. *Review of Economics and Statistics*, 103(5):857–873.

Correia, S., Guimarães, P., and Zylkin, T. (2019). ppmlhdfe: Fast Poisson estimation with high-dimensional fixed effects. *Working Paper*.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2):445–510.

Dhyne, E., Kikkawa, A. K., Mogstad, M., and Tintelnot, F. (2021). Trade and domestic production networks. *Review of Economic Studies*, 88(2):643–668.

Eaton, J., Kortum, S. S., and Kramarz, F. (2022). Firm-to-firm trade: Imports, exports, and the labor market. *Working Paper*.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the Penn World Table. *American Economic Review*, 105(10):3150–82.

Ghosh, A. (2023). Religious divisions and production technology: Experimental evidence from India. *Working Paper*.

Goraya, S. S. (2023). How does caste affect entrepreneurship? Birth versus worth. *Journal of Monetary Economics*, 135:116–133.

Guiso, L., Sapienza, P., and Zingales, L. (2009). Cultural biases in economic exchange? *Quarterly Journal of Economics*, 124(3):1095–1131.

Gupta, B. (2014). Discrimination or social networks? industrial investment in colonial India. *Journal of Economic History*, 74(1):141–168.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, pages 153–161.

Helpman, E., Melitz, M., and Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics*, 123(2):441–487.

Hnatkovska, V., Hou, C., and Lahiri, A. (2021). Convergence across castes. *Working Paper*.

Hnatkovska, V., Lahiri, A., and Paul, S. (2012). Castes and labor mobility. *American Economic Journal: Applied Economics*, 4(2):274–307.

Hnatkovska, V., Lahiri, A., and Paul, S. B. (2013). Breaking the caste barrier: Intergenerational mobility in india. *Journal of Human Resources*, 48(2):435–473.

Huneeus, F. (2018). Production network dynamics and the propagation of shocks. *Working Paper*.

India Justice Report (2023). *India Justice Report: Ranking States on Police, Judiciary,* *Prisons and Legal Aid.*

Kaufman, D. and Kraay, A. (2023). *Worldwide Governance Indicators.*

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86.

Lim, K. (2018). Endogenous production networks and the business cycle. *Working Paper.*

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review*, 111(6):1807–1844.

Macchiavello, R. and Morjaria, A. (2015). The value of relationships: Evidence from a supply shock to Kenyan rose exports. *American Economic Review*, 105(9):2911–45.

Mitra, A. and Ray, D. (2014). Implications of an economic theory of conflict: Hindu-Muslim violence in India. *Journal of Political Economy*, 122(4):719–765.

Munshi, K. (2014). Community networks and the process of development. *Journal of Economic Perspectives*, 28(4):49–76.

Munshi, K. (2019). Caste and the Indian economy. *Journal of Economic Literature*, 57(4):781–834.

Munshi, K. and Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1):46–98.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *Quarterly Journal of Economics*, 122(2):569–600.

Oberfield, E. (2018). A theory of input–output architecture. *Econometrica*, 86(2):559–589.

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools. *American Economic Review*, 109(3):774–809.

Rao, M. (2023). Front-line courts as state capacity: Evidence from India. *Working Paper.*

Rauch, J. E. (1996). Trade and search: Social capital, sogo shosha, and spillovers. *Working Paper.*

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of international Economics*, 48(1):7–35.

Rauch, J. E. and Trindade, V. (2002). Ethnic Chinese networks in international trade. *Review of Economics and Statistics*, 84(1):116–130.

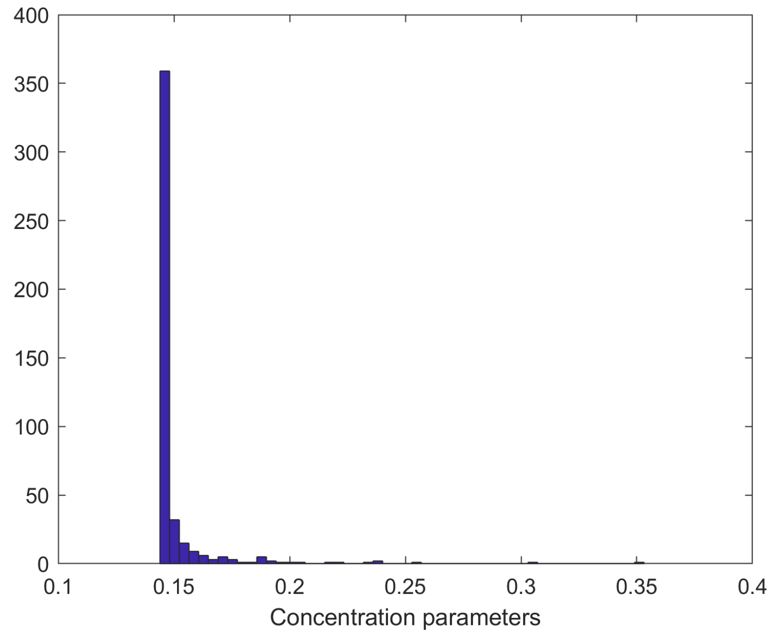
Schoar, A., Iyer, R. K., and Kumar, S. (2008). Importance of ethnic networks for business transactions of the small enterprises. *Working Paper.*

Startz, M. (2016). The value of face-to-face: Search and contracting problems in Nigerian trade. *Working Paper.*

Taschereau-Dumouchel, M. (2019). Cascades and fluctuations in an economy with an endogenous production network. *Working Paper.*

A Additional figures and tables

Figure A1: Histogram of estimated concentration parameters for Dirichlet distribution



Notes: Estimated concentration parameters for a Dirichlet distribution according to the maximum likelihood estimation from Equation 18.

Table A1: Estimation for matching cost, second stage

	(1)	(2)
BC	0.165***	0.143***
	(0.010)	(0.010)
$BC \times \mathbb{I}^{court}$		0.121***
		(0.018)
ihs $\left[\widehat{n(z, z')}\right]$	6.091***	6.097***
	(0.038)	(0.038)
Obs.	5,628,290	5,626,102
Pseudo R2	0.159	0.160

Notes: Column 1 shows the results of estimating Equation 20. Column 2 adds an interaction with $\mathbb{I}^{court}$, which indicates if the sum of the delays in the origin-district court and the destination-district court is above the 75th percentile. We winsorize $\ln \widehat{n(z, z')}$ 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 in parentheses. The higher the Bhattacharyya coefficient, the culturally closer two firms are.

B Targeted and untargeted moments

Outdegree and indegree

Data. In our dataset, for each firm i, we calculate the number of firms it sold to and the number of firms it bought from. These are the outdegree and the indegree, respectively. Then, we divide these number by the total number of firms in our sample to normalize them.

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 follow a similar procedure to calculate the number of sellers each buyer z' has.

Intermediate sales and prices

Data. In our dataset, for each firm i, we calculate the sales to other firms. We normalize this measure by dividing this measure by the total sales. For the prices, we take the log unit price of each good traded at the HS 6-digit level.

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 total sales. For the prices, we recover the elements of price matrix $p\left(z, z'\right)$.

Table A2: Targeted and untargeted moments

Targeted Moments		
	Data	Model
Log outdegree, average	-9.24	-9.48
Log outdegree, variance	0.98	0.86
Log intermediate sales, variance	2.82	2.83
Log indegree, average	-9.39	-9.16
Untargeted Moments		
	Data	Model
Log indegree, variance	0.61	0.17
Log intermediate prices, variance	5.77	3.93

C Additional specifications

C.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.[22] 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 $\varepsilon\rightarrow 0^{+}$ such that KL_{sym} is well-defined. Tables A3 and A4 show the regression results for the intensive margin, extensive margin and unit prices. 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.

[22] This interpretation diverts from the standard use the Kullback-Leibler has in information theory, where a higher divergence means a higher information loss.

Table A3: Effect of cultural proximity on trade, intensive and extensive margins, Kullback-Leibler

	(1)	(2)	(3)
Dep. Variable	Log Sales	Log Transactions	Trade Indicator
KL_{sym}	-0.005***	-0.003**	-0.00004***
	(0.002)	(0.001)	(0.00000)
Obs.	32,843	32,843	5,628,290
Adj. R2	0.410	0.356	0.0106
FE	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.

Notes: Columns 1 and 2 show the results of estimating a modified version of Equation 1. Column 3 shows the result of estimating a modified version of Equation 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 et al., 2019).

Table A4: Effect of cultural proximity on prices, Kullback-Leibler

	(1)	(2)	(3)
Dep. Variable	Log Prices	Log Prices	Log Prices
KL_{sym}	0.002*	0.002**	0.001
	(0.001)	(0.001)	(0.001)
Obs.	235,001	236,617	230,900
Adj. R2	0.933	0.925	0.936
FE	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 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 et al., 2019).

C.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.[23] We do not find evidence that hierarchies (and preference-based discrimination) across social groups matter for our cultural proximity results.

We 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 indicates the seller is placed below the buyer in the social hierarchy. We include these two indicators by interacting them with our measure of cultural proximity. Table A5 presents the results for the intensive and extensive margins. The baseline category is that both firms belong to the same hierarchy. First, we find the baseline coefficient is very similar to those of Table 1. 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.

[23] While the Varna-based hierarchy only relates to the Hindu religion, we also place other religions in this hierarchy based on their income levels.

Table A5: Effect of cultural proximity on trade by vertical hierarchies, intensive and extensive margins

	(1)	(2)	(3)
Dep. Variable	Log Sales	Log Transactions	Trade Indicator
BC	0.129***	0.079***	0.0010***
	(0.035)	(0.029)	(0.0001)
$BC \times \mathbb{I}_{\nu_H \omega_L}$	0.008	0.072	-0.0003
	(0.116)	(0.092)	(0.0003)
$BC \times \mathbb{I}_{\nu_L \omega_H}$	-0.027	-0.123	-0.0004
	(0.129)	(0.103)	(0.0002)
Obs.	31,119	31,119	5,477,548
Adj. R2	0.412	0.357	0.0107
FE	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.

Notes: Columns 1 and 2 show the results of estimating a modified version of Equation 1. Column 3 shows the result of estimating a modified version of Equation 2. ***, ** and * indicate statistical significance at the 99, 95, and 90 percent levels 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 the two firms are. The number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia et al., 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.

C.3 Goods specialization

The cultural groups in India are, in many cases, defined by the production of specific goods (Munshi, 2019).[24] In this section, we study if the reason behind the cultural proximity results are 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 transaction level, we estimate the regression

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

$$\ln n_g(\nu,\omega,t) = \iota_{\nu g} + \iota_{gt} + \iota_{\omega} + \iota_{od}(\nu,\omega) + \delta BC(\nu,\omega) + \xi\left(BC(\nu,\omega)\times \mathbb{I}_g^{spec}\right) + \epsilon_g(\nu,\omega), \quad \text{(A1)}$$

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 A6 presents the results for the sales. The results suggest that specialization does not play a role in the determination of the effect of cultural proximity on trade.

Table A6: Effect of cultural proximity on trade by good specialization, intensive margin

	(1)	(2)
Dep. Variable	Log Sales	Log Sales
BC	0.064***	0.064***
	(0.023)	(0.023)
$BC \times \mathbb{I}_g^{spec,seller}$	0.135	
	(0.304)	
$BC \times \mathbb{I}_g^{spec,buyer}$		0.185
		(0.118)
Obs.	229,719	229,719
Adj. R2	0.854	0.854
FE	Seller×HS, buyer, month×HS, origin×dest.	Seller×HS, buyer, month×HS, origin×dest.

Notes: This table shows the results of estimating Equation A1. ***, ** and * indicate statistical significance at the 99, 95 and 90 percent level respectively. Good g is defined according to 6-digit HS classification. Sales were 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 the two firms are. The number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia et al., 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.

C.4 Number of varieties sold and bought

In this part, we analyse whether the cultural proximity results prevail for firms that sell and buy more varieties of goods. We first count how many 4-digit HS codes a firm buys or sells. Table A7 presents the results for the intensive margin, following a modified version of Equation 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 that they trust. Moreover, this explanation based on trust is compatible with the results related to differentiated goods from Section 3.1. In both cases, we posit that cultural proximity acts as a coping mechanism to market imperfections stemming from low-quality institutions in India.

Table A7: 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 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.

C.5 Age of firms

In this section, we follow Becker (1957) to analyse whether taste-based discrimination is behind our main findings. If there is taste-based discrimination, then we should see that firms that sell to culturally closer firms at lower prices are willing to forego profits because of their preferences. A consequence would be that these firms are more prone to go bankrupt.

For our empirical analysis, we leverage information on the establishment date from IndiaMART (the date on which a firm was established). If there is taste-based discrimination, then we should see older firms relying less on cultural proximity. This would mean that firms that had a preference for selling to firms culturally closer to them eventually went bankrupt, while the survivors were those firms that did not show these preferences.

Table A8 shows the results for a modified version of the intensive margin regressions according to Equation 1. If there was taste-based discrimination, then the interaction between the measure of cultural proximity and age should have a negative coefficient. However, we find weak evidence for taste-based discrimination, such that we cannot conclude this is the reason behind our results.

Table A8: Effect of cultural proximity after controlling for establishment age of sellers (IndiaMART), intensive margin

	(1)	(2)
Dep. Variable	Log Sales	Log Transactions
BC	0.800**	0.479
	(0.371)	(0.311)
$BC \times$Log age seller	-0.207*	-0.122
	(0.112)	(0.091)
Obs.	5,859	5,859
Adj. R2	0.387	0.237
FE	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.

Notes: This table shows the results of estimating a modified version of Equation 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. Number of observations varies between specifications due to the dropping of observations separated by a fixed effect (Correia et al., 2019). Age of seller comes from data reported in IndiaMART.

C.6 Industry fixed effects

In this section, we revise the intensive margin regressions after considering the fact that there are pairs of industries that are bound to trade more than other pairs. For instance, perhaps certain castes happen to specialize in certain industries, and these industries are more likely to trade with each other. For this analysis, we add an industry of seller × industry of buyer fixed effect to Equation 1. The sectors are based on the 4-digit HS code of the good with the highest sales for each firm. Table A9 presents the results. When compared to the results in Table 1, we find that, while the effect of cultural proximity is slightly higher, the main result prevails.

Table A9: Effect of cultural proximity after controlling for industries, intensive margin

	(1)	(2)
Dep. Variable	Log Sales	Log Transactions
BC	0.145***	0.104**
	(0.055)	(0.045)
Obs.	16,229	16,229
Adj. R2	0.395	0.308
FE	Seller, buyer, origin×dest., seller ind.×buyer ind.	Seller, buyer, origin×dest., seller ind.×buyer ind.

Notes: This table shows the results of estimating Equation 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. Industry classified according to the 4-digit HS classification of the most sold good by each firm. 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 et al., 2019).

C.7 Cancellations

Our data is unique in that it records canceled transactions as well. Among the diverse reasons for which cultural proximity could affect trade, we can also study reneged contracts. In this section, we analyse whether it becomes more likely pairs of firms will cancel their transactions if they are far in cultural terms. We estimate the following specification:

$$\mathbb{I}_{g}^{cancel}\left(\nu,\omega,t\right)=\iota_{\nu g}+\iota_{gt}+\iota_{\omega}+\iota_{od}\left(\nu,\omega\right)+\delta BC\left(\nu,\omega\right)+\epsilon_{g}\left(\nu,\omega,t\right), \tag{A2}$$

where $\mathbb{I}_{g}^{cancel}\left(\nu,\omega,t\right)$ is a dummy that says if there was at least one canceled transaction going from firm ν to firm ω of good g (at the 6-digit HS classification) in month t, $\iota_{\nu g}$ is a seller-good fixed effect, ι_{gt} is a good-month fixed effect and ι_{ω} is a seller fixed effect. Here, we control for the month of the year to account for macro events that could have caused widespread cancellations.

Table A10 presents the results. We find that the closer firms are in cultural terms, the less likely it is that there will be a cancellation. Here we must highlight that cancellations can occur for reasons other than reneged contracts. Therefore, our results are just suggestive of there being a channel whereby cultural proximity reflects in lower reneged contract, but are not conclusive.

Table A10: Cancellations

	(1)	(2)	(3)
Dep. Variable	Ever canceled (0/1)		
BC	-0.005*	-0.000	-0.005*
	(0.003)	(0.002)	(0.003)
Obs.	256,819	258,481	252,829
Adj. R2	0.102	0.0695	0.108
FE	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 A2 at the transaction level. ***, ** 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 et al., 2019).

C.8 Correction for selection bias

Helpman et al. (2008) suggest that the traditional gravity equation estimations, which we use for our intensive margin regressions, are biased because of selection issues. Thus, in this section, we follow their proposed correction for selection bias.

In the first stage, we estimate the following linear probability model:

$$tr\left(\nu,\omega\right)=\iota_{\nu}+\iota_{\omega}+\iota_{od}\left(\nu,\omega\right)+\delta BC\left(\nu,\omega\right)+\gamma B2B\left(\nu,\omega\right)+\epsilon\left(\nu,\omega\right), \tag{A3}$$

where we follow the nomenclature and the in-state sample from our extensive margin regressions. Here, we need an excluded instrument that affects only the extensive margin (i.e. the matching cost) and not the intensive margin (i.e. the trade cost). Thus, we consider the indicator variable $B2B\left(\nu,\omega\right)$ that equals 1 when both seller ν and buyer ω are in IndiaMART and equals 0 otherwise. As mentioned in Section 2.3, IndiaMART is the largest e-commerce platform for business-to-business (B2B) transactions in India.[25] Thus, the idea here is that it is easier for both firms to match if they take part in this platform.

We present the results of this first stage in Table A11. As before, the closer the firms are in cultural terms, the more likely it is they will trade. Additionally, if both firms participate in IndiaMART, the more likely the trade.

After the estimation, we recover the predicted probability of trading $\widehat{tr}\left(\nu,\omega\right)$, with which we calculate the latent variable

$$\widehat{\zeta}\left(\nu,\omega\right)=\Phi^{-1}\left(\widehat{tr}\left(\nu,\omega\right)\right),$$

where $\Phi^{-1}\left(\cdot\right)$ is the inverse CDF of the standard normal distribution.

Following Heckman (1979), we obtain the inverse Mills ratio

$$\Upsilon\left(\widehat{\zeta}\right)=\frac{\phi\left(\widehat{\zeta}\left(\nu,\omega\right)\right)}{\Phi\left(\widehat{\zeta}\left(\nu,\omega\right)\right)},$$

where $\phi\left(\cdot\right)$ is the PDF of the standard normal distribution, and $\Phi\left(\cdot\right)$ is the CDF of the standard normal distribution.

[25] In 2019, there were between 5 and 6 million registered firms in IndiaMART (https://www.indiamart.com), which represented all firm size groups and all geographic regions in India.

For the second stage, we estimate

$$\ln y(\nu,\omega) = \iota_{\nu} + \iota_{\omega} + \iota_{od}(\nu,\omega) + \delta BC(\nu,\omega) + \upsilon\Upsilon\left(\widehat{\zeta}\right) + \epsilon(\nu,\omega), \tag{A4}$$

where $y(\nu,\omega)$ is the total positive sales of seller ν to buyer ω and the term $\Upsilon\left(\widehat{\zeta}\right)$ accounts for selection bias.

We present the second stage results in Columns 3 and 4 of Table A12. We must note that, because of computational reasons, we work with only an in-state sample such that our results are not directly comparable to the baseline results from Table 1. Therefore, in Columns 1 and 2 of Table A12 we present the results with the in-state sample but without the correction for selection bias.

We find that not considering the correction for selection biases the coefficients downwards. This way, we conclude that the main results related to the intensive margin in the paper represent a lower bound of the effect of cultural proximity on trade.

Table A11: Correction for selection bias, first stage

Dep. Variable	Trade Dummy
BC	0.0010***
	(0.0001)
$B2B$	0.0016***
	(0.0003)
Obs.	5,628,290
Adj. R2	0.0106
FE	Seller, buyer, origin×dest.

Notes: Table shows the results of estimating Equation A3. Sample contains only in-state firms. ***, ** 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. The higher the Bhattacharyya coefficient, the culturally closer two firms are.

Table A12: Correction for selection bias, second stage

	(1)	(2)	(3)	(4)
Dep. Variable	Log Sales	Log Transactions	Log Sales	Log Transactions
BC	0.148***	0.095***	0.223***	0.132**
	(0.035)	(0.029)	(0.074)	(0.055)
Inv. Mills rat.			0.503	0.246
			(0.421)	(0.298)
Obs.	26,238	26,238	26,238	26,238
Adj. R2	0.392	0.360	0.392	0.360
FE	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.	Seller, buyer, origin×dest.

Notes: Columns 1, 2, 3 and 4 show the results of estimating Equation A4. Columns 1 and 2 do not consider the correction for selection bias term. Sample contains only in-state firms. ***, ** 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. 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 et al., 2019).

D Model derivations

In this section, we include details about the derivations of the theoretical model.

D.1 Technology

A unique variety ω is produced by a single firm which minimizes its unit cost of production subject to its production technology, so

$$\min_{\{m(\nu,\omega)\}} \quad \int_{\nu\in\Omega(\omega)} m\left(\nu,\omega\right) p\left(\nu,\omega\right) d\nu + wl\left(\omega\right), s.t.$$

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

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

$$y\left(\omega\right) = 1.$$

Merge the first and third constraints, such that

$$\begin{aligned} y(\omega) &= \kappa_\alpha z(\omega) l(\omega)^\alpha m(\omega)^{1-\alpha}, \\ 1 &= \kappa_\alpha z(\omega) l(\omega)^\alpha m(\omega)^{1-\alpha}, \\ l(\omega)^\alpha &= \frac{1}{\kappa_\alpha z(\omega) m(\omega)^{1-\alpha}}, \\ &= \kappa_\alpha^{-1} z(\omega)^{-1} m(\omega)^{\alpha-1}, \\ l(\omega) &= \kappa_\alpha^{-\frac{1}{\alpha}} z(\omega)^{-\frac{1}{\alpha}} m(\omega)^{\frac{\alpha-1}{\alpha}}. \end{aligned}$$

Rewrite the minimization problem, such that

$$\begin{aligned} &\min_{\{m(\nu,\omega)\}} \int_{\nu\in\Omega(\omega)} m(\nu,\omega) p(\nu,\omega) d\nu + wl(\omega), \\ &\int_{\nu\in\Omega(\omega)} m(\nu,\omega) p(\nu,\omega) d\nu + \kappa_\alpha^{-\frac{1}{\alpha}} wz(\omega)^{-\frac{1}{\alpha}} m(\omega)^{\frac{\alpha-1}{\alpha}}, \\ &\int_{\nu\in\Omega(\omega)} m(\nu,\omega) p(\nu,\omega) d\nu + \kappa_\alpha^{-\frac{1}{\alpha}} wz(\omega)^{-\frac{1}{\alpha}} \left(\int_{\nu\in\Omega(\omega)} m(\nu,\omega)^{\frac{\sigma-1}{\sigma}} d\nu\right)^{\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}}. \end{aligned}$$

The first order condition with respect to $m(\nu,\omega)$ is

$$\begin{aligned} 0 &= p(\nu,\omega) + \kappa_\alpha^{-\frac{1}{\alpha}} wz(\omega)^{-\frac{1}{\alpha}} \left(\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}\right)(\dots)^{\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1}\left(\frac{\sigma-1}{\sigma}\right) m(\nu,\omega)^{\frac{\sigma-1}{\sigma}-1}, \\ p(\nu,\omega) &= \kappa_\alpha^{-\frac{1}{\alpha}}\left(\frac{1-\alpha}{\alpha}\right) wz(\omega)^{-\frac{1}{\alpha}}(\dots)^{\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1} m(\nu,\omega)^{-\frac{1}{\sigma}}, \\ m(\nu,\omega)^{\frac{1}{\sigma}} &= \frac{\kappa_\alpha^{-\frac{1}{\alpha}}\left(\frac{1-\alpha}{\alpha}\right) wz(\omega)^{-\frac{1}{\alpha}}(\dots)^{\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1}}{p(\nu,\omega)}, \\ m(\nu,\omega) &= \frac{\kappa_\alpha^{-\frac{\sigma}{\alpha}}\left(\frac{1-\alpha}{\alpha}\right)^\sigma w^\sigma z(\omega)^{-\frac{\sigma}{\alpha}}(\dots)^{\sigma\left(\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1\right)}}{p(\nu,\omega)^\sigma}. \end{aligned}$$

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

$$
\begin{aligned}
\frac{m\left(\nu,\omega\right)}{m\left(\nu',\omega\right)} &= \frac{\frac{\kappa_{\alpha}^{-\frac{\sigma}{\alpha}}\left(\frac{1-\alpha}{\alpha}\right)^{\sigma} w^{\sigma} z(\omega)^{-\frac{\sigma}{\alpha}} (\dots)^{\sigma\left(\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1\right)}}{p(\nu,\omega)^{\sigma}}}{\frac{\kappa_{\alpha}^{-\frac{\sigma}{\alpha}}\left(\frac{1-\alpha}{\alpha}\right)^{\sigma} w^{\sigma} z(\omega)^{-\frac{\sigma}{\alpha}} (\dots)^{\sigma\left(\frac{\sigma}{\sigma-1}\frac{\alpha-1}{\alpha}-1\right)}}{p(\nu',\omega)^{\sigma}}}, \\
&= \frac{\frac{z(\omega)^{-\frac{\sigma}{\alpha}}}{p(\nu,\omega)^{\sigma}}}{\frac{z(\omega)^{-\frac{\sigma}{\alpha}}}{p(\nu',\omega)^{\sigma}}}, \\
&= \frac{p\left(\nu',\omega\right)^{\sigma}}{p\left(\nu,\omega\right)^{\sigma}}, \\
m\left(\nu',\omega\right) &= \frac{p\left(\nu,\omega\right)^{\sigma} m\left(\nu,\omega\right)}{p\left(\nu',\omega\right)^{\sigma}}.
\end{aligned}
$$

We plug this expression back into the expression for the composite of intermediates, so

$$
\begin{aligned}
m\left(\omega\right) &= \left(\int_{\nu'\in\Omega(\omega)} m\left(\nu',\omega\right)^{\frac{\sigma-1}{\sigma}} d\nu\right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\int_{\nu'\in\Omega(\omega)} \left(\frac{p\left(\nu,\omega\right)^{\sigma} m\left(\nu,\omega\right)}{p\left(\nu',\omega\right)^{\sigma}}\right)^{\frac{\sigma-1}{\sigma}} d\nu\right)^{\frac{\sigma}{\sigma-1}}, \\
&= p\left(\nu,\omega\right)^{\sigma} m\left(\nu,\omega\right) \underbrace{\left(\int_{\nu'\in\Omega(\omega)} p\left(\nu',\omega\right)^{1-\sigma} d\nu\right)^{\frac{\sigma}{\sigma-1}}}_{=\left(P(\omega)^{1-\sigma}\right)^{\frac{\sigma}{\sigma-1}}}, \\
&= p\left(\nu,\omega\right)^{\sigma} m\left(\nu,\omega\right) \left(P\left(\omega\right)^{1-\sigma}\right)^{\frac{\sigma}{\sigma-1}}, \\
&= p\left(\nu,\omega\right)^{\sigma} m\left(\nu,\omega\right) P\left(\omega\right)^{-\sigma}, \\
&= m\left(\omega\right) p\left(\nu,\omega\right)^{-\sigma} P\left(\omega\right)^{\sigma}, \\
p\left(\nu,\omega\right) m\left(\nu,\omega\right) &= m\left(\omega\right) p\left(\nu,\omega\right)^{1-\sigma} P\left(\omega\right)^{\sigma}, \\
n\left(\nu,\omega\right) &= P\left(\omega\right) m\left(\omega\right) p\left(\nu,\omega\right)^{1-\sigma} P\left(\omega\right)^{\sigma-1}, \\
&= N\left(\omega\right) p\left(\nu,\omega\right)^{1-\sigma} P\left(\omega\right)^{\sigma-1},
\end{aligned}
$$

which is the demand of firm ω from variety ν, where $P\left(\omega\right)^{1-\sigma} = \int_{\nu\in\Omega(\omega)} p\left(\nu,\omega\right)^{1-\sigma} d\nu$ is the price index faced by firm ω, $n\left(\nu,\omega\right) = p\left(\nu,\omega\right) m\left(\nu,\omega\right)$ is the expenditure of ω on variety ν, and $N\left(\omega\right) = P\left(\omega\right) m\left(\omega\right)$ is the total expenditure of firm ω.

The expression for the 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

$$\begin{aligned} \max_{\{p(\nu,\omega)\}} \quad \int_{\omega \in \Omega(\nu)} & (p(\nu,\omega) - d(\nu,\omega)\, c(\nu))\, m(\nu,\omega)\,, s.t. \\ & m(\nu,\omega) = m(\omega)\, p(\nu,\omega)^{-\sigma} P(\omega)^{\sigma}\,, \end{aligned}$$

where $d(\nu,\omega)$ is the iceberg cost of firm ν selling to ω. Rewrite the profit function $\pi(\nu,\omega)$, such that

$$\begin{aligned} \pi(\nu,\omega) &= (p(\nu,\omega) - d(\nu,\omega)\, c(\nu))\, m(\nu,\omega)\,, \\ &= p(\nu,\omega)\, m(\nu,\omega) - d(\nu,\omega)\, c(\nu)\, m(\nu,\omega)\,, \\ &= p(\nu,\omega)\, m(\omega)\, p(\nu,\omega)^{-\sigma} P(\omega)^{\sigma} - d(\nu,\omega)\, c(\nu)\, m(\omega)\, p(\nu,\omega)^{-\sigma} P(\omega)^{\sigma}\,, \\ &= m(\omega)\, p(\nu,\omega)^{1-\sigma} P(\omega)^{\sigma} - d(\nu,\omega)\, c(\nu)\, m(\omega)\, p(\nu,\omega)^{-\sigma} P(\omega)^{\sigma}\,. \end{aligned}$$

The first order condition is

$$\begin{aligned} [p(\nu,\omega)] : (1-\sigma)\, m(\omega)\, p(\nu,\omega)^{-\sigma} P(\omega)^{\sigma} & \\ - (-\sigma)\, d(\nu,\omega)\, c(\nu)\, m(\omega)\, p(\nu,\omega)^{-\sigma-1} P(\omega)^{\sigma} &= 0, \\ (\sigma-1)\, m(\omega)\, p(\nu,\omega)^{-\sigma} P(\omega)^{\sigma} &= \sigma d(\nu,\omega)\, c(\nu)\, m(\omega)\, p(\nu,\omega)^{-\sigma-1} P(\omega)^{\sigma}\,, \\ (\sigma-1) &= \sigma d(\nu,\omega)\, c(\nu)\, p(\nu,\omega)^{-1}\,, \\ p(\nu,\omega) &= \left(\frac{\sigma}{\sigma-1}\right) c(\nu)\, d(\nu,\omega)\,, \\ &= \mu c(\nu)\, d(\nu,\omega)\,, \end{aligned}$$

where $\mu = \frac{\sigma}{\sigma-1}$ is the markup.

D.2 Preferences

A representative household maximizes its utility subject to its budget constraint, so

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

The first order condition with respect to firm ω is

$$\begin{aligned}
&[y\left(\omega\right)] : \left(\frac{\sigma}{\sigma-1}\right)\left(\dots\right)^{\frac{\sigma}{\sigma-1}-1}\left(\frac{\sigma-1}{\sigma}\right) y\left(\omega\right)^{\frac{\sigma-1}{\sigma}-1} = \lambda P\left(\omega\right),\\
&\lambda P\left(\omega\right) = \left(\dots\right)^{\frac{\sigma}{\sigma-1}-1} y\left(\omega\right)^{-\frac{1}{\sigma}},
\end{aligned}$$

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

$$\lambda P\left(\omega'\right) = \left(\dots\right)^{\frac{\sigma}{\sigma-1}-1} y\left(\omega'\right)^{-\frac{1}{\sigma}}.$$

We then divide both first-order conditions, such that

$$\begin{aligned}
\frac{\lambda P\left(\omega\right)}{\lambda P\left(\omega'\right)} &= \frac{\left(\dots\right)^{\frac{\sigma}{\sigma-1}-1} y\left(\omega\right)^{-\frac{1}{\sigma}}}{\left(\dots\right)^{\frac{\sigma}{\sigma-1}-1} y\left(\omega'\right)^{-\frac{1}{\sigma}}},\\
\frac{P\left(\omega\right)}{P\left(\omega'\right)} &= \frac{y\left(\omega\right)^{-\frac{1}{\sigma}}}{y\left(\omega'\right)^{-\frac{1}{\sigma}}},\\
&= \frac{y\left(\omega'\right)^{\frac{1}{\sigma}}}{y\left(\omega\right)^{\frac{1}{\sigma}}},\\
y\left(\omega'\right)^{\frac{1}{\sigma}} &= y\left(\omega\right)^{\frac{1}{\sigma}}\, \frac{P\left(\omega\right)}{P\left(\omega'\right)},\\
y\left(\omega'\right) &= y\left(\omega\right)\left(\frac{P\left(\omega\right)}{P\left(\omega'\right)}\right)^{\sigma}.
\end{aligned}$$

We plug this demand back into 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 a variety ω.

D.3 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.

D.4 Equilibrium given network

In this section, we derive the expression for the equilibrium objects given the structure of the production network. We first derive the recursive expression for prices, and then for total sales.

Recursive expression for prices. Consider the expression for the CES price index, so

$$
\begin{aligned}
P(\omega)^{1-\sigma} &= \int_{\nu\in\Omega(\omega)} p(\nu,\omega)^{1-\sigma}\, d\nu, \\
P(z')^{1-\sigma} &= \int p(z,z')^{1-\sigma}\, l(z,z')\, dG(z), \\
&= \int \left(\left(\frac{\sigma}{\sigma-1}\right) c(z)\, d(z,z')\right)^{1-\sigma} l(z,z')\, dG(z), \\
&= \mu^{1-\sigma} \int \left(c(z)\, d(z,z')\right)^{1-\sigma} dG(z), \\
&= \mu^{1-\sigma} \int \left(\frac{P(z)^{1-\alpha}}{z} d(z,z')\right)^{1-\sigma} l(z,z')\, dG(z), \\
&= \mu^{1-\sigma} \int \left(\frac{P(z)^{1-\alpha}}{z} d(z,z')\right)^{1-\sigma} l(z,z')\, dG(z), \\
&= \mu^{1-\sigma} \int P(z)^{(1-\alpha)(1-\sigma)}\, z^{\sigma-1} d(z,z')^{1-\sigma}\, l(z,z')\, dG(z).
\end{aligned}
$$

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.