NBER WORKING PAPER SERIES

BOUND BY ANCESTORS:
IMMIGRATION, CREDIT FRICTIONS, AND GLOBAL SUPPLY CHAIN FORMATION

Jaerim Choi
Jay Hyun
Ziho Park

Working Paper 31157
http://www.nber.org/papers/w31157

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2023

Park is grateful for financial support from the National Science and Technology Council of Taiwan. We are grateful to JaeBin Ahn, Pol Antras, Costas Arkolakis, Dany Bahar, Xue Bai, Samuel Bazzi (discussant), Matilde Bombardini, Barthelemy Bonadio, Leah Boustan, Miguel Cardoso, Yongsung Chang, Jaedo Choi, Jeonghwan Choi, Hyunbae Chun, Andrew Dickens, Rob Feenstra, Hisham Foad, Stefania Garetto, Jean-Francois Gauthier, Raphael Godefroy, Linda Goldberg, Robin Gong, Juan Carlos Hallak, Tarek Hassan, Federico Huneeus, Hanwei Huang, Jung Hur, Xian Jiang, Gaurav Khanna, Daisoon Kim, Gueyon Kim, Se-jik Kim, Wookun Kim, Eunjee Kwon, Yoonsoo Lee, Inessa Love, Caroline Le Pennec (discussant), Emile Marin, Chris Meissner, Camelia Minoiu, Jakob Munch, Nicolas Morales, Giovanni Peri, Kadee Russ, Tim Schmidt-Eisenlohr, Shu Shen, Yang Shen, Yongseok Shin, Ina Simonovska, Guillaume Sublet, Chang Sun, Deborah Swenson, Mathieu Taschereau-Dumouchel, Alan Taylor, Martin Uribe, Jonathan Vogel (discussant), Shang-Jin Wei (discussant), and seminar participants in Academia Sinica, Applied Economics Workshop, Brock University, FREIT-EIIT (UCSD), Hanyang University, HEC Montréal, Hong Kong Trade Seminar, KIET, Louisiana State University, Migration and Organizations Conference (Wharton), Midwest International Trade Conference (Notre Dame), Midwest Macroeconomics Meeting (SMU), NBER Conference (Central Bank of Chile), Seoul National University, Sogang University, UC Davis, UC Santa Cruz, Université de Montréal, and Yonsei University for invaluable comments and suggestions. We thank Pol Antras, Sofia Bauducco, Linda Goldberg, and Sebnem Kalemli-Ozcan for inviting this work at the the NBER Conference on International Fragmentation, Supply Chains, and Financial Frictions, hosted by the Central Bank of Chile, and Shang-Jin Wei (discussant) for invaluable discussion and suggestions during the conference. First draft: June 20, 2022. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Bound by Ancestors: Immigration, Credit Frictions, and Global Supply Chain Formation
Jaerim Choi, Jay Hyun, and Ziho Park
NBER Working Paper No. 31157
April 2023
JEL No. F14,F22,F36,F60,G30,J61,L14

ABSTRACT

This paper shows that the ancestry composition shaped by century-long immigration to the US can explain the current structure of global supply chain networks. Using an instrumental variable strategy, combined with a novel dataset that links firm-to-firm global supply chain information with a US establishment database and historical migration data, we find that the co-ethnic networks formed by immigration have a positive causal impact on global supply chain relationships between foreign countries and US counties. Such a positive impact not only exists in conventional supplier-customer relationships but also extends to strategic partnerships and trade in services. Examining the causal mechanisms, we find that the positive impact is stronger for counties in which more credit-constrained firms are located and that such a stronger effect becomes even more pronounced for foreign firms located in countries with weak contract enforcement. Collectively, the results suggest that co-ethnic networks serve as social collateral to overcome credit constraints and facilitate global supply chain formation.

Jaerim Choi
Department of Economics
University of Hawaii at Manoa
Saunders Hall 542
2424 Maile Way
Honolulu, HI 96822
choijm@hawaii.edu

Ziho Park
Department of Economics
National Taiwan University
No. 1, Section 4, Roosevelt Rd
Da’an District, Taipei City, 10617
Taiwan
zihopark@ntu.edu.tw

Jay Hyun
Department of Applied Economics
HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal, Québ H3T 2A7
Canada
jungsik.hyun@hec.ca

1 Introduction

During the past decade, the global economy has experienced a series of events that have triggered disturbances in global supply chains and immigration. For instance, the reorganization of global supply chains has been at the center of economic and political debates following the 2018 US-China trade war, the COVID-19 pandemic, and the Russo-Ukrainian War, which have all severely disrupted global supply chain networks. Despite decades of movements toward global integration through global value chains, the recent COVID-19 pandemic, for example, has ignited a movement toward "renationalization" under the premise that global integration has exacerbated the negative impacts of the pandemic.[1] Just as importantly, these events have also triggered disruptions in immigration across borders. For example, the Trump administration tightened visa access for high-skilled foreign workers in 2020 as part of its protectionist policies, and the COVID-19 pandemic led many countries to implement various types of border restrictions and halt immigration processes.

The increased attention to global supply chain reorganization following the massive disruption in global supply chains and immigration then raises the following questions. What determines the global supply chain network structure? Does immigration, among its potential determinants, affect the formation of global supply chains across countries? If so, through what mechanisms does immigration shape the global supply chain structure? In this article, we address these questions by establishing a causal linkage between immigration and global supply chain formation: Immigration from a given country to a given US county has a positive causal impact on firm-to-firm global supply chain relationships between the two regions.[2] In terms of the underlying economic mechanisms, we provide novel evidence that co-ethnic networks across borders play an important role in mitigating credit constraint problems, which fosters global supply chain formation.

Understanding the role of immigration in the formation of global supply chain networks is contemporaneously relevant because if immigration affects global supply chain formation, then it implies that the recent decline in immigration due to the pandemic could have long-lasting impacts on the global supply chain structure above and beyond the direct impact of the pandemic. Moreover, the investigation of the interdependence between two seemingly unrelated networks—co-ethnic and global supply chain networks—and their interaction with credit frictions can deepen our understanding of the complexities of global supply chains, including credit provision across borders, in the era of the

[1] As reported by the United Nations Department of Economic and Social Affairs in April 2020, "*The adverse effects of prolonged restrictions on economic activities in developed economies will soon spill over to developing countries via trade and investment channels. ⋯ In addition, global manufacturing production could contract significantly amid the possibility of extended disruptions to global supply chains.*"— World Economic Situation and Prospects: April 2020 Briefing, No. 136.

[2] Researchers have typically considered the structure of supply chains or production networks as exogenously given. A few recent studies (Lim, 2018; Acemoglu and Azar, 2020; Kopytov et al., 2022; Taschereau-Dumouchel, 2022) develop a tractable framework to investigate the *endogenous* determination of the production network, but mostly focus on *domestic* firm-to-firm networks. Our work is differentiated from these studies, as we empirically explore the determination of *global* firm-to-firm supply chain networks.

globalization of trade and finance with the rising interplay between the two.

To this end, we first explore whether the co-ethnic networks between US counties and foreign countries, which are formed by century-long immigration from foreign countries to the US, have *causally* shaped the current landscape of the global supply chain networks between these locations. However, establishing the causal relationship is empirically challenging because it requires exogenous variation in ethnic composition that is unrelated to global firm-to-firm supply chains. Certain characteristics, such as climate similarity between two places, may induce more immigrants from a given country to reside in a particular area and, at the same time, create more global supply chain networks between the two locations. Furthermore, past supply chain networks between an origin country and a destination county may have led to increased immigration.

We solve these issues by adopting a "leave-out push-pull" approach, an instrumental variable technique pioneered by Burchardi et al. (2019), which exploits the unique US migration history for more than a century to construct an instrument for the present distribution of ancestry composition. The US has experienced successive waves of immigration from many different origin countries, with immigrants settling in different destinations within the US depending on their relative attractiveness in each period. Three factors, (i) a push factor, (ii) a pull factor, and (iii) a recursive factor, summarize the time-series variation in the relative attractiveness of different destinations within the US combined with the staggered arrival of migrants from different origins. The push factor measures the total number of migrants arriving in the US from origin country o in time t, the pull factor captures the relative economic attractiveness of destination d to migrants arriving in time t, and the recursive factor gauges the relative size of the pre-existing local population of the ancestry of origin country o in destination d at time t. The instrumental variable is constructed as the full set of the simple and higher-order interactions of these pull and push factors.

Equipped with this empirical strategy, we assemble a unique and comprehensive dataset that links US firms' global supply chain relationships from the FactSet Revere with their establishment-level information from the National Establishment Time-Series (NETS) database. We then relate these global supply chain relationships to migration network data (i.e., century-long data on historical migration from foreign countries to US counties) and find a positive impact of co-ethnic networks on the current structure of global supply chain networks. Quantitatively, doubling the number of residents with ancestry from a given origin country relative to the sample mean (from 320 to 640) increases by 4.9 (or 4.5) percentage points the probability that at least one firm in destination d has a supplier (or customer) headquartered in origin country o. We also find that the positive impact of co-ethnic networks on global supply chain linkages operates at the intensive margin: the number of suppliers and customers.

We conduct various robustness checks to corroborate the key empirical findings. We show that global supply chain formation is a phenomenon that is distinct from foreign direct investment (FDI):

Controlling for FDI and strictly ruling out any potential within-multinational linkages do not alter our results. We conduct a placebo test and find that the impact of ancestry composition on global supply chain linkages works only through trade-engaging establishments—which directly engage in firm-to-firm trade with foreign companies—and not through establishments that do not engage in trade. The core results are also robust to restricting firms to single-establishment firms or defining firms' US county locations based on their headquarters locations. Additionally, neither using ancestry compositions prior to 2010 (i.e., 1980, 1990, and 2000) nor restricting supply chain linkages to those newly formed after 2010 affects our main results, reassuring that the reverse causality is not driving our core results. Our results are also robust to dropping Asian countries and US counties on the West Coast, implying that the results are not particularly driven by the strong tie between these regions. We also confirm that dropping top ancestry origins or top global supply chain partner countries do not alter our results. Finally, the results are robust to using alternative measure of ancestry composition.

We further explore whether the positive impact of co-ethnic networks goes beyond the conventional supplier-customer relationships in the literature. First, using a strategic relationship variable (e.g., joint ventures or research collaboration) in FactSet Revere, we uncover that co-ethnic networks have a positive impact on a broader form of relationships—strategic partnerships. Next, we find that the effect of ancestry composition on global supply chain formation is positive for both manufacturing and non-manufacturing, with the effects being slightly stronger for non-manufacturing establishments. Collectively, the first key finding of our paper provides evidence that co-ethnic networks play an important role in shaping the current global supply chain networks, and that such an effect operates mainly through trade-engaging establishments and extends further to non-conventional supply chain relationships such as strategic partners and non-manufacturing industries.

Finally, we explore the mechanisms responsible for the key findings. It has been well documented that business across borders, compared to within-border business, entails considerable barriers such as the poor enforcement of international contracts, incomplete information about foreign partners, and cultural and language differences. In these frictional environments, Rauch (2001) noted that co-ethnic networks might help mitigate these informal barriers. Therefore, we posit that the role of co-ethnic networks in forming global supply chain relationships becomes more important under circumstances where credit and contractual friction issues are more serious.[3,4]

[3] Because two or more unrelated firms engage in firm-to-firm global supply chains, unlike environments in which firms can internalize the issue of financial problems, firms rely heavily on informal finance, such as trade credit, to address the mismatch between payment and delivery. It has been noted that trust plays a major role in granting and receiving trade credit, especially when there are no formal contractual enforcement mechanisms in place because implicit financing contracts naturally involve default risks. Moreover, such trust is enhanced by social networks, which facilitate informal borrowing (Karlan et al., 2009). We argue that a co-ethnic network is one such social network that enables each involved party to build trust in trade credit relationships.

[4] An example of informal financing that crucially depends on social trust—especially among immigrants of the same co-ethnicity—is rotating credit associations (RCAs). For example, Korean immigrants who settled in Los Angeles have dealt with financial constraints through RCAs in doing business with vendors in South Korea. We describe such

Using firms' trade credit performance information, we validate the hypothesis that the positive impact of co-ethnic networks on global supply chain relationships becomes stronger for US counties in which more credit-constrained firms are located. Furthermore, using a triple interaction among the co-ethnic networks, credit constraints, and judicial quality of origin countries, we further show that such a stronger effect becomes even more pronounced for origin countries with weak contract enforcement. These results, taken together, suggest that co-ethnic networks serve as a form of social collateral to overcome credit constraints and facilitate global supply chain formation.

Related Literature

Our paper studies the interplay of immigration, global supply chains, and credit constraints under incomplete contract enforcement. Gould (1994) and Head and Ries (1998) are two early examples of studies using a gravity model to uncover the positive impact of immigration on trade for the US and Canada, respectively. More recent studies have attempted to establish a more rigorous causal effect of immigration on international trade (Rauch and Trindade, 2002; Aleksynska and Peri, 2014; Cohen et al., 2017; Parsons and Vézina, 2018; Cardoso and Ramanaryanan, 2019; Bonadio, 2020) and FDI (Javorcik et al., 2011; Burchardi et al., 2019). Our paper contributes to the literature by providing a causal linkage between immigration and firm-to-firm global supply chain relationships and by examining the underlying economic mechanism.

The distinction between firm-to-firm supply chain networks and within-firm activities such as FDI has been a central topic in the theory of the firm. For instance, as the seminal work of Coase (1937) emphasized, within-firm activities are governed by organizational rules, whereas firm-to-firm transactions are mediated by prices (i.e., market mechanisms). Therefore, issues involving prices, trade credit, and trade under incomplete contracting are more pronounced in firm-to-firm transactions than they are in the within-firm movement of factors (Williamson, 1975; Hart and Moore, 1988).[5] By focusing on global supply chain relationships, we can explore credit constraints and incomplete contracting environments, which are not directly applicable to FDI (e.g., Burchardi et al., 2019).

This paper also complements the literature on the nexus between social networks and credit constraints under incomplete contracting. It has been noted in the sociology literature that social networks play an essential role in building trust in relationships (e.g., Putnam et al., 2000). Following this insight, in the economics literature, Karlan et al. (2009) developed a theory of trust based on informal contract enforcement in social networks. They found that social networks create trust when agents use connections as social collateral to facilitate informal borrowing.[6] As such, social networks

an example in more detail in Section 5.2.

[5]In particular, this distinction matters more in international trade due to the difficulties of formal contractual enforcement in international transactions. As a result, a domestic firm faces an important choice problem of whether to trade with a separate foreign firm or to own a subsidiary in the foreign country, and this choice problem has been an active research agenda in the international trade literature (Antras and Helpman, 2004).

[6]In Karlan et al. (2009), they noted that "*the possibility of losing valuable friendships secures informal transactions*

can build trust that serves as a form of social collateral to mitigate credit constraints (Besley et al., 1993; Fafchamps, 2000; Karlan, 2005; Wu et al., 2014; Levine et al., 2018).[7] Although our credit friction mechanism is in the spirit of these papers, we find this channel in the reciprocal relationship between immigration and global supply chains, which has not been previously investigated. Furthermore, consistent with Antras and Foley (2015), we find that the role of social networks in alleviating credit constraints is more pronounced for foreign countries with weaker contract enforcement.

Finally, our work relates to the literature on the role of financial frictions in international transactions. At the national level, it has been well established that financial development serves as a source of comparative advantage (Kletzer and Bardhan, 1987; Beck, 2002; Matsuyama, 2005; Ju and Wei, 2010, 2011; Manova, 2013). At the firm level, credit-constrained firms exhibit worse performance in terms of trade participation (Manova et al., 2015; Chaney, 2016), and financial shocks can result in a decline in firm-level trade activities (Amiti and Weinstein, 2011; Niepmann and Schmidt-Eisenlohr, 2017). As a consequence, multinationals and FDI may help overcome such credit constraints in international transactions (Desai et al., 2004; Antras et al., 2009; Manova et al., 2015) by internalizing financial frictions within firms, and multinationals fare better during financial crises at least partially for this reason (Alfaro and Chen, 2012). From a different angle, we contribute to the literature on the nexus between trade and finance by focusing on firm-to-firm international transactions instead of multinationals by showing that the co-ethnic networks have the potential to alleviate credit frictions and facilitate global *firm-to-firm* transactions.

The remainder of this paper is organized as follows. Section 2 introduces the datasets for our main analysis. Section 3 presents the estimation model. Section 4 presents the main results. Section 5 provides evidence on the credit constraint mechanism. Section 7 concludes.

2 Data

We construct a unique dataset of US firms' global supply chain relationships merged with their establishment-level information. The data on these global supply chain networks come from the FactSet Revere database, with which we are able to identify customer-supplier relationships between US and foreign firms. The US firms in FactSet Revere might have multiple establishments (i.e., multiple locations) across the US. The National Establishment Time-Series (NETS) database contains the precise location information on all establishments, which enables us to identify global firm-to-firm supply chain relationships between US county d and foreign country o. These global supply chain relationships are then linked to migration networks between US county d and origin country o,

in the same way that the possibility of losing physical collateral can secure formal lending."

[7]Korean rotating credit associations (RCAs) in the US are good examples of trust-based social networks that facilitate the entrepreneurship of immigrants by reducing the severity of financial obstacles (Light et al., 1990). Moreover, using Chinese firm-level data, Kong et al. (2020) find that the hometown connections of CEOs, one example of social network connections, increase access to trade credit.

wherein the immigration and ancestry data come directly from Burchardi et al. (2019).

2.1 Global Supply Chain Relationships

The FactSet Revere database is designed to uncover business relationship interconnections among companies globally. FactSet analysts systematically collect companies' relationship information from primary public sources such as SEC 10-K annual filings, investor presentations, and press releases. They provide four normalized relationship types: (i) customers, (ii) suppliers, (iii) competitors, and (iv) strategic partners. We identify a firm's supplier and customer firms using relationship types categorized as "suppliers" and "customers."

The FactSet Revere supply chain relationship database covers approximately 200,000 firms, including more than 30,000 publicly listed firms around the world, comprising over 725,000 unique business relationships, with historical data going back as far as 2003.[8] Importantly, the FactSet Revere database includes both publicly listed and private firms, provides both important and less important relationships, and incorporates relationship information obtained from both the direct disclosure of a source company and the reverse disclosure of another company regarding the source company.

These features allow researchers to better capture the comprehensive picture of the network structure in the economy, providing an important advantage compared to other firm-level supply chain data sources used in US studies such as the Compustat Segment database, which primarily relies on SEC 10-K filings that require publicly listed firms to disclose those customers that account for more than 10% of the firms' revenue.[9] Additionally, to give a concrete example of the importance of reverse disclosure, the FactSet Revere Data and Methodology Guide states that there are 22 suppliers directly disclosed by an exemplary firm—Walmart—while, as a result of the reverse relationships methodology, an additional 293 companies disclose Walmart as a customer, therefore providing a more extensive supplier network for the exemplary firm. Finally, a linkage weight between a supplier and a customer is disclosed whenever available, measured as the percentage of the supplier's revenue arising from its relationship with that customer.

2.2 National Establishment Time-Series Database

The National Establishment Time-Series (NETS) database is an annual panel of a near universe of US establishments. The NETS tracks firm structures systematically, assigning each establishment

[8] Company information is fully reviewed annually, and changes based on corporate actions are monitored daily. Thus, it provides a detailed and up-to-date dataset, including information on inter-company relationships. The coverage statistics are based on the authors' own calculations.

[9] Therefore, the Compustat Segment database captures only those important relationships directly disclosed by publicly listed firms. In contrast, the Factset Revere database collects information not only from SEC 10-K filings but also from many other sources, such as investor presentations and press releases, and incorporates reverse disclosure by other firms, allowing researchers to capture less important relationships and relationships that involve private firms.

a unique identifier called the Data Universal Numbering System (DUNS) number, which remains unchanged even after mergers and acquisitions (M&As), and is not reused after an establishment exit. Each establishment is also assigned an identifier of its headquarters establishment. Using this feature of the NETS, we group together all the establishments under the same firm, and match them to the corresponding firm in FactSet Revere.

Importantly for our purpose, the NETS includes establishment-level information about an exact address and a precise location at the longitude and latitude levels. Since the NETS links each establishment to its headquarters establishment, we can potentially identify multiple locations of a firm that can be matched to the FactSet Reserve database.

The location information in the NETS is obtained from the Yellow Pages, public records, and other wide-ranging data verification efforts. A number of studies have demonstrated the accuracy of location information in the NETS. For example, Barnatchez et al. (2017) find that the county-level correlation between the NETS and the Census Bureau's County Business Patterns (CBP) is above 0.99 regarding both employment and establishment counts. Based on the documented indication for data accuracy, a number of recent studies have utilized the location information in the NETS as their main dataset (e.g., Gray et al., 2015; Rossi-Hansberg et al., 2021; Behrens et al., 2022; Hyun and Kim, 2022; Hyun et al., 2022; Oberfield et al., 2022).

Importantly, for each establishment, the NETS data provide information on yearly trade status. Specifically, we have information on whether an establishment engages in export, import (or both) in a given year. Using this information, we define *trade-engaging* establishments as those who engage in trade activities in a given year. Since supply chain information is only defined at the firm level, not the establishment level, for a given supply chain linkage between a foreign firm and a US firm, we connect the foreign country with US counties by using location information of trade-engaging establishments of the US firm.[10]

The source of the NETS data is Dun & Bradstreet (D&B), which, as a credit rating company, has a strong incentive to collect accurate data. Using reports from firms' trade credit suppliers, D&B calculates the PayDex score, a 100-point credit score measure for the overall manner of trade credit payment, which is widely used as a measure of creditworthiness (e.g., Kallberg and Udell, 2003; Borisov et al., 2021; Avramidis et al., 2022). We use PayDex scores—both the yearly maximum and minimum PayDex scores—to investigate whether co-ethnic networks help mitigate credit constraints in global supply chains.

[10]The underlying premise is that for a given US firm, the firm's trade-engaging establishments are more likely to directly engage in firm-to-firm trade activities compared to its non-trade-engaging establishments. We use location of non-trade-engaging establishments to perform placebo tests in Section 4.2.2.

2.3 Immigration and Ancestry

The immigration and ancestry data come directly from Burchardi et al. (2019) where they used the individual files of the Integrated Public Use Microdata Series (IPUMS) samples of the 1880, 1900, 1910, 1920, 1930, 1970, 1980, 1990, and 2000 waves of the US census and the 2006-2010 five-year sample of the American Community Survey. For each wave, the largest available sample is chosen; observations are weighted by personal weights to obtain a representative sample. The individual-level data are then aggregated to the level of historic US counties and foreign countries and are then transformed into the 1990 country-county level using various transition matrices.[11]

For our empirical analysis, the flow of immigrants, $I^t_{o,d}$, is measured as the number of migrants from foreign origin country o to US destination county d between $t-1$ and t (the interval between two consecutive census waves).[12] The stock of ancestry information, $A^t_{o,d}$, is measured as the number of respondents in US county d who report the first ancestry as foreign origin country o at time t. The dyadic dataset covers 3,141 US counties, 195 foreign countries, and 10 census waves.

In the baseline analyses, we exploit the first 9 waves (1880-2000) of migrations and exclude the 2000-2010 wave, following Burchardi et al. (2019).[13] We show, however, that the inclusion of the 10th wave barely affects our main result.

2.4 Other Datasets

We use geographic distances, absolute latitude differences, and measures of agricultural similarity between origin country o and US county d as a set of control variables.[14] The distance variables are measured as the distance between the coordinates for all postal codes within a US county and the coordinates of the main city in a foreign country.[15] The agricultural similarity between each origin country o and US county d is measured as the difference in the crop suitability of the country and county for a select group of crops.[16]

We also use an industry-level measure of external finance dependency from Rajan and Zingales (1998). Specifically, this widely used measure captures an industry's need for external finance, which is defined as the difference between investments and the internal cash generated from operations, originally sourced from data on U.S. firms. In addition to the PayDex score measures, this industry-

[11] For a more detailed procedure, please refer to Burchardi et al. (2019).

[12] In the first sample, i.e., the 1880 census, the variable, $I^{1880}_{o,d}$, is defined as the number of residents who were either born in o or whose parents were born in o.

[13] As is well documented in Jensen et al. (2015) and Burchardi et al. (2019), the number of foreign-born first-generation immigrants in the 2010 American Community Survey is somewhat understated. Therefore, we exclude the 2000-2010 wave in the baseline regression.

[14] The three measures come directly from Burchardi et al. (2019).

[15] If a US county has multiple postcodes, then one of them is randomly selected. The associated geo-coordinates can be downloaded from GeoNames (geonames.org) and CEPII (cepii.fr).

[16] The raw data come from the Food and Agriculture Organization of the United Nations Global Agro-Ecological Zones (FAOGAEZ).

level measure complements the analysis in investigating the credit constraint channel.

Finally, to gauge the extent to which countries enforce contracts pursuant to an agreement, we use a measure of judicial quality from Nunn (2007), which is originally drawn from the "rule of law" from Kaufmann et al. (2004). This judicial quality measure enables us to further examine the credit constraint mechanism by exploiting institutional differences across countries.

2.5 Summary Statistics

Table 1 provides the basic summary statistics of the datasets, for four different levels of the unit: (i) origin-destination level, (ii) origin level, (iii) destination level, and (iv) firm level.

In Panel A, one notable pattern in the global supply chain relationships (the outcome variables of our interest) is that only 3% of the origin-destination pairs have at least one linkage between foreign suppliers (in origin) and US customers (in destination) in 2011-2014. A similar pattern holds for foreign-customer-and-US-supplier linkages, meaning that most of the variation in global supply chain relationships is along the extensive margin. Along the same line, conditional on having at least one foreign-supplier-and-US-customer linkage, the average number of foreign suppliers within an origin-destination pair is 2.54, with a standard deviation of 4.27. We find a similar pattern for foreign-customer-and-US-supplier linkages.

Turning our attention to the key independent variable (i.e., the ancestry variable) in Panel A, the average number of residents with ancestry from a particular origin country is 320 in 2010. The standard deviation of the ancestry variable is 5,960, which provides us with sufficient variation for our empirical analysis.

Additionally, Panel A provides credit-related variables and a measure of judicial quality, which are converted into the origin-destination level.[17] The PayDex score measures display sufficient variations across origin-destination pairs, which enables us to investigate the role of credit constraints in exploring mechanisms. The measure of judicial quality, which is converted from the origin level to the origin-destination level, also has sufficient variation across observations with a mean of 0.50 and a standard deviation of 0.21.

Panels B and C of Table 1 provide the summary statistics of origin countries and destination counties, respectively. In Panel B (origin level), the average origin country is connected to 105 US counties by having suppliers headquartered in the origin country who supply to US firms in those counties. Similarly, the average origin country is connected to 118 US counties through customer linkages. The distribution is highly skewed, which is also confirmed by Figure A.1. Panel C summarizes similar information at the destination level. We find a highly skewed distribution of the number of origins connected through global supply chains, which is visualized in Figure A.2.

[17] D&B provides PayDex score information at the establishment level. We average the scores of the establishments within a county. After this procedure, at the origin-destination level (the unit of analysis), the number of observations shrinks to 592,995 due to some non-reporting observations.

Table 1: Summary Statistics

Panel A	Origin-Destination-Level					
Variable	Obs.	Mean	Std. Dev.	P10	P50	P90
I(N.Supp>0)	612495	0.03	0.18	0.00	0.00	0.00
I(N.Cust>0)	612495	0.04	0.19	0.00	0.00	0.00
N.Supp (>0)	20385	2.54	4.27	1.00	1.00	5.00
N.Cust (>0)	22968	2.44	4.29	1.00	1.00	5.00
Ancestry 2010 (in thousands)	612495	0.32	5.96	0.00	0.00	0.08
Distance (km)	612495	9122.39	3802.10	3358.40	9330.90	14027.00
Latitude Difference (degree)	612495	19.44	11.31	3.92	19.53	34.52
100-PayDexMax	592995	24.65	3.53	20.57	24.36	28.47
100-PayDexMix	592995	30.26	5.62	23.15	30.15	36.13
Judicial Quality (JQ)	452304	0.50	0.21	0.29	0.45	0.86
Panel B	**Origin-Level**					
Variable	Obs.	Mean	Std. Dev.	P10	P50	P90
N.Destination - Supplier Linkage	195	104.54	213.18	0.00	0.00	448.00
N.Destination - Customer Linkage	195	117.78	223.95	0.00	1.00	506.00
Avg. Ancestry 2010 (in thousands) across Destination	195	0.32	1.30	0.00	0.01	0.48
Avg. Distance (km) across County	195	9122.39	3711.71	3658.20	9279.65	14006.68
Avg. Latitude Difference (degree) across Destination	195	19.44	10.19	5.39	20.29	33.79
Judicial Quality (JQ)	144	0.50	0.21	0.29	0.45	0.86
Panel C	**Destination-Level**					
Variable	Obs.	Mean	Std. Dev.	P10	P50	P90
N.Origin - Supplier Linkage	3141	6.49	11.56	0.00	0.00	24.00
N.Origin - Customer Linkage	3141	7.31	12.92	0.00	0.00	26.00
Avg. Ancestry 2010 (in thousands) across Origin	3141	0.32	1.13	0.01	0.07	0.63
Avg. Distance (km) across Origin	3141	9122.39	443.46	8592.80	9062.89	9722.44
Avg. Latitude Difference (degree) across Origin	3141	19.44	3.28	15.71	19.08	23.55
N.Establishment	3141	219.64	701.33	9.00	46.00	437.00
N.Establishment (Trade-engaging)	3141	9.99	46.41	0.00	1.00	16.00
100-PayDexMax	3103	24.76	3.74	20.57	24.40	28.73
100-PayDexMix	3103	30.49	6.05	23.15	30.25	36.59
Panel D	**Firm-Level**					
Variable	Obs.	Mean	Std. Dev.	P10	P50	P90
N.Supp (>0)	3306	2.83	3.86	1.00	1.00	6.00
N.Cust (>0)	2954	3.97	5.13	1.00	2.00	11.00
N.Establishment	48458	14.35	169.98	1.00	2.00	13.00
N.Establishment (Trade-engaging)	48458	0.66	3.38	0.00	0.00	1.00
100-PayDexMax	36813	25.28	8.81	20.00	22.24	33.45
100-PayDexMix	36813	31.39	12.59	20.00	28.15	46.00

Notes: This table presents the number of observations (Obs), mean, standard deviation, and tenth (P10), fiftieth (P50), and ninetieth (P90) percentiles of the variables. Panel A refers to our main sample of (origin) country and (destination) county pairs. Panels B and C refer to the origin country and destination county samples, respectively. Panel D refers to the sample of firms in US counties. I(N.Supp>0) is a dummy variable that equals one if any firm whose trade-engaging establishments located in destination county d has at least one supplier firm headquartered in origin country o between 2011 and 2014, and zero otherwise. I(N.Cust>0) is similarly defined, where we use customer firms headquartered in origin o instead of supplier firms. N.Destination - Supplier (Customer) Linkage in Panel B indicates the number of destination counties for each origin country, where destination counties have firms whose supplier (customer) firms are headquartered in origin o. N.Origin - Supplier (Customer) Linkage in Panel C indicates the number of origin countries for each destination county, whose supplier (customer) firms are headquartered in origin countries. The averaged variables (Avg.) in Panels B and C are simple averages of the respective variables in Panel A.

Panel C additionally provides the summary statistics of the number of establishments as well as the number of establishments that engage in trade between 2011 and 2014. On average, there are 220 establishments, among which 10 establishments engage in trade. This is consistent with the well-established fact in the literature that only a few businesses engage in foreign transactions (e.g., Bernard et al., 2007).

Finally, Panel D shows the summary statistics at the firm level. The total number of firms in our sample is 48,458 and an average firm has approximately 14 establishments. However, an average firm has less than one (0.66) trade-engaging establishment, which is again consistent with the well-established fact in trade. Among 48,458 firms, 3,306 (2,954) firms supply to (buy from) at least one foreign company.

2.6 Descriptive Analyses

Before laying out our empirical specification and moving to formal regression analyses, we examine a general correlation between the number of ancestry and the number of global supply chain (GSC) linkages.

Panel A of Table B.1 ranks the top thirty origin countries in terms of the number of ancestry in the United States in 2010. These countries account for 92.9% of the total ancestry groups. The largest economies in Europe (such as Germany and the UK), countries that share borders with the US (Mexico and Canada), and populous Asian countries (such as China and India) are on this list. Similarly, Panel B ranks the top thirty countries in terms of the global supply chain linkages between 2011 and 2014. These countries account for 88.2% of the global supply chain linkages in this period. Notably, 17 countries are on both Panels, which suggests a positive connection between the two lists.

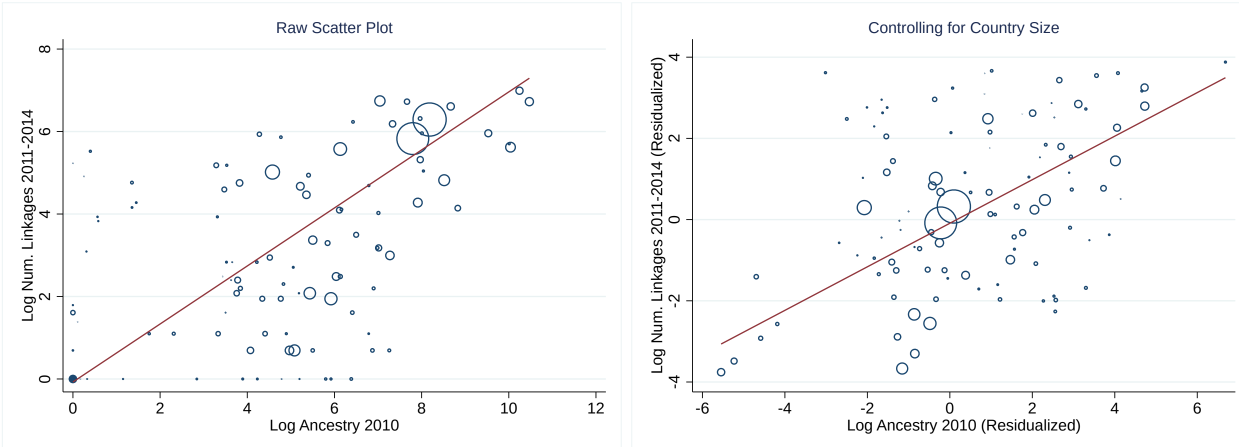
The positive association between these two panels in Table B.1 is further confirmed in the left panel of Figure 1, which shows a clear positive relationship between the number of ancestry and the number of global supply chain linkages at the origin-country level. That is, a country from which more immigrants originated and moved to the US tends to have more global supply chain linkages with the United States. Such a relationship prevails even after controlling for country size in the right panel of Figure 1.[18] In Table B.2 of Appendix B, we further show that such a positive correlation remains after controlling for other country characteristics, such as GDP per capita and distance to the US, in addition to country population.[19]

Moreover, for each major origin country, we document a positive correlation between the number of ancestry in 2010 and the number of global supply chain linkages from 2011 to 2014 at the US

[18] Specifically, we use country population to proxy for country size. In the right panel of Figure 1, both variables on the y-axis and the x-axis are *residualized*, obtained by conditioning on the country population. We rely on the Frisch-Waugh theorem to calculate the residualized log ancestry and log number of linkages.

[19] In addition to country population, we control for GDP per capita, distance and latitude difference with US, the measure of ethnic fractionalization (Alesina et al., 2003), FDI indicator, and continent fixed effects. Note that, by controlling for population and GDP per capita, we are effectively controlling for country-level GDP.

Figure 1: Global Supply Chain Linkages and Ancestry at the Origin Country Level



Notes. This figure plots the relationship between the number of ancestry and the number of global supply chain linkages at the origin country level. The left panel shows the scatter plot and its linear fit from the raw data; the right panel shows the scatter plot and its linear fit after controlling for country size (measured by country population). Specifically, we first apply the Frisch-Waugh theorem to partial out the role of country size and obtain the *residualized* log ancestry and log number of linkages. Then, we plot the residualized number of ancestry against the residualized number of linkages. The size of the circle reflects the size of the origin countries (measured by population).

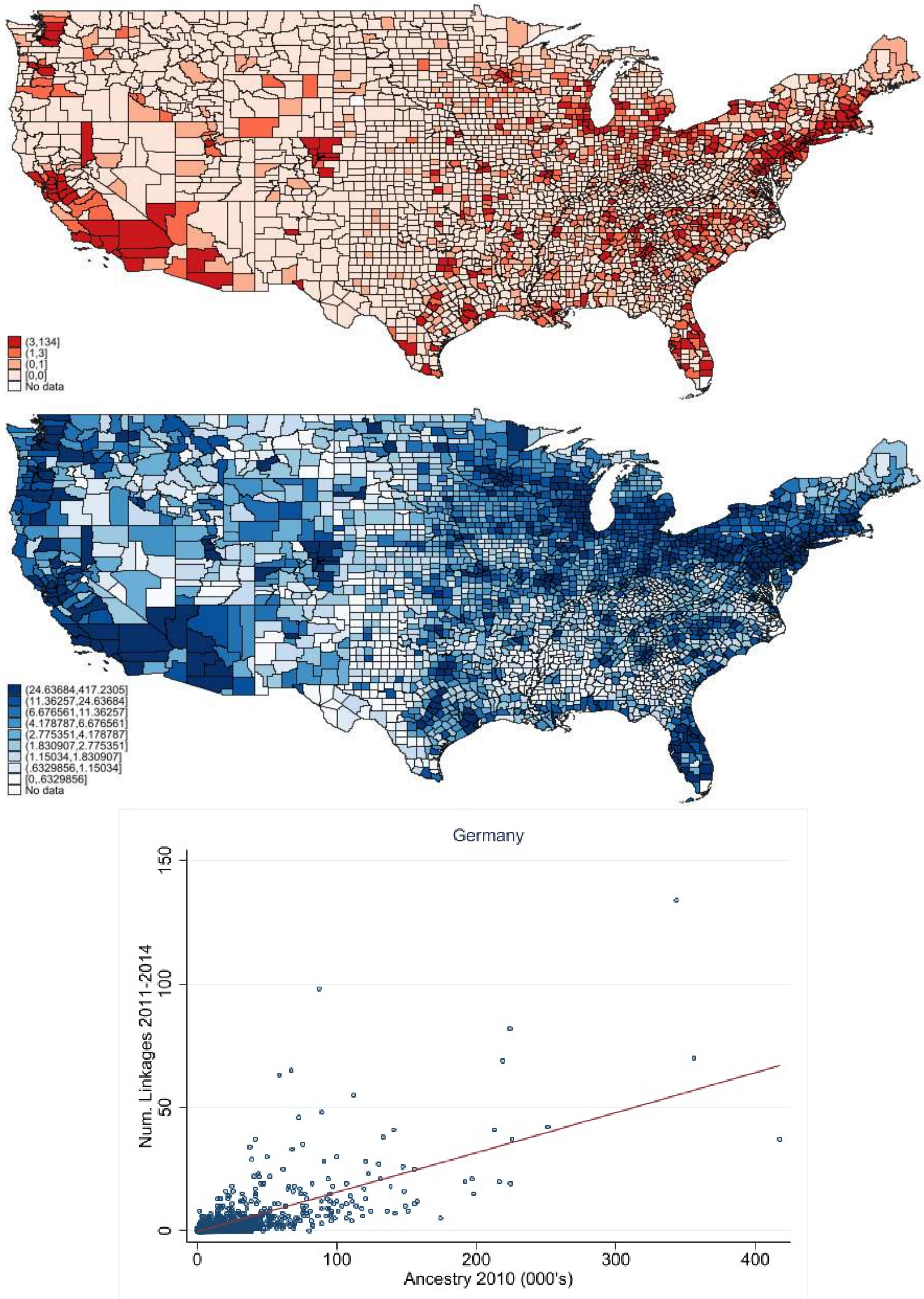
county level. First, in Figure 2, we restrict our sample to Germany, which is the country with the largest number of ancestry in the United States in 2010. The red and blue maps display global supply chain linkages with German firms and German ancestry in the US, respectively, both at the county level. These two maps have a clear overlap, which suggests a link between the two county-level variations. This positive relationship is also confirmed in the scatter plot between the number of German ancestry and the number of global supply chain linkages with German firms, where each dot represents a US county. From Figures A.3 to A.12 in Appendix A, we document similar patterns for some of the countries that were identified as major origin countries in Panel A of Table B.1. The same positive associations persist for other top five countries (UK, Mexico, Ireland, and Italy), Canada, and top five Asian countries on the list (China, Philippines, India, South Korea, and Japan).

In the next section, we move on to a formal regression analysis to establish a causal linkage between ancestry composition and global supply chain linkages.

3 Empirical Strategy

Our aim is to estimate the causal impact of immigration on global supply chain relationships. Specifically, we would like to evaluate whether the presence of descendants of migrants from an origin country o within US county d induces more US firms in county d to have a global supply chain link (either as a supplier or a customer) with foreign firms based in the origin country o. To do so, we set

Figure 2: Global Supply Chain Linkages and Ancestry: Germany



Notes. The figure is restricted to one origin country: Germany. The first red map shows the county variation in the global supply chain linkages with German firms within the US. The second blue map shows the county variation in German ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of German ancestry and the number of global supply chain linkages with German firms across US counties.

up the following structural gravity equation:

$$Y_{o,d} = \delta_o + \delta_d + \beta A_{o,d}^{2010} + X'_{o,d}\gamma + \varepsilon_{o,d} \tag{3.1}$$

where $Y_{o,d}$ is the outcome of interest regarding the global supply chain relationships between origin o and destination d. We construct both a dummy variable and the number of linkages to explore extensive and intensive margins, respectively, and consider supplier and customer linkages separately. Specifically, I(N.Supp>0) is a dummy variable that equals one if any firm whose trade-engaging establishments located in destination county d has at least one supplier firm headquartered in origin country o between 2011 and 2014, and zero otherwise.[20] I(N.Cust>0) is similarly defined, where we use customer firms headquartered in origin o instead of supplier firms. We also consider the log of the number of linkages to explore the intensive margin, using a logarithm of the number of supplier firms (Log N.Supp) and the number of customer firms (Log N.Cust).

$A_{o,d}^{2010}$ is defined as the log of one plus the number of residents in US county d that report having ancestors in origin country o in 2010. $X'_{o,d}$ is a set of control variables that include geographic distance, differences in latitude, and agricultural similarity between o and d. δ_o are origin country fixed effects; δ_d are destination county fixed effects. Standard errors are clustered at the origin-country level. The key coefficient of our interest, β, captures the causal impact of immigration on a global supply chain relationship between firms in o and d.

The standard OLS estimation applied to Equation (3.1) will be biased if the migration network term $A_{o,d}^{2010}$ and the error term $\varepsilon_{o,d}$ are correlated, even after controlling for origin and destination fixed effects and a set of control variables. For instance, origin-destination-specific omitted factors, such as climate affinity or other characteristics, can simultaneously affect global supply chain networks today as well as historical migration flows. Alternatively, perhaps the direction of causality can be the opposite such that past global supply chain relationships might drive migration.

To address these challenges, we adopt a "leave-out push-pull" approach (a variant of the IV strategy), originally developed by Burchardi et al. (2019), which enables us to identify the causal impact of immigration on global supply chain networks. In summary, the "leave-out push-pull" approach is based on a simple dynamic model of migrations that comprise three factors: (i) a push factor, (ii) a pull factor, and (iii) a recursive factor. Migrations from origin country o to US county d in period t, $I_{o,d}^t$, depend on the total number of migrants arriving in the US from o in t (I_o^t: push factor), the relative economic attractiveness of d to migrants arriving in t ($\frac{I_d^t}{I^t}$: pull factor), and the relative size of the pre-existing local population of ancestry o in d at t ($\frac{A_{o,d}^{t-1}}{A_o^{t-1}}$: recursive factor).

[20] Trade-engaging establishments are defined using NETS data, which provide information on trade activity for each establishment. We define establishments to be engaged in trade if they report being active in trade at least once between 2011 and 2014.

Following Burchardi et al. (2019), the first-stage specification can be expressed as follows:[21]

$$A_{o,d}^{2010} = \delta_o + \delta_d + \sum_{t=1880}^{2000} \alpha_t I_{o,-r(d)}^t \frac{I_{-c(o),d}^t}{I_{-c(o)}^t} + \sum_{n=1}^{5} \delta_n PC_n + X_{o,d}'\gamma + \eta_{o,d} \tag{3.2}$$

where $I_{o,-r(d)}^t$ captures the push factor, excluding migrants to d's census division; $\frac{I_{-c(o),d}^t}{I_{-c(o)}^t}$ measures the pull factor, excluding migrants from the continent on which origin country o is located. The exclusion of migrants to d's census division and migrants from o's continent further ensures that $I_{o,d}^t$ and $\varepsilon_{o,d}$ are uncorrelated. PC_n denotes the n-th principal component summarizing the information contained in the higher-order terms $I_{o,-r(d)}^s \cdots I_{o,-r(d)}^t \frac{I_{-c(o),d}^t}{I_{-c(o)}^t}, \forall t < s \leq 2010$.

The key identifying assumption is that conditional on the control variables and fixed effects, $I_{o,-r(d)}^t \frac{I_{-c(o),d}^t}{I_{-c(o)}^t}$ and $\varepsilon_{o,d}$ are uncorrelated. In the empirical procedures, we further relax this assumption and include the origin x destination's census division fixed effects ($\delta_o \times \delta_{r(d)}$) and destination x continent-of-origin fixed effects ($\delta_d \times \delta_{c(o)}$). Those fixed effects control for any confounding factors that operate within origin-by-census-division and destination-by-continent pairs, respectively. Specifically, origin x destination's census division fixed effects allow us to compare the impact of immigration from *the same origin country* to US counties *within the same census division*, thereby controlling for the role of country size as well as any census-division-wise factors that attract immigrants from the origin. Similarly, destination x continent-of-origin fixed effects allow us to compare the impact of immigration to *the same destination county* originating from countries *within the same continent*. Thus, these fixed effects absorb any continent-wise preferences toward a given destination county, as well as any destination county-specific factors such as county amenities or county's weather condition.

4 Main Results

This section shows that the ancestry composition shaped by immigrants to the US for more than a century can causally explain the current structure of global supply chain networks. We first present the instrumental variable regression results and then provide various robustness checks to corroborate our findings.

4.1 Instrumental Variable Regressions

Table 2 presents the estimated results of Equation (3.1) using IV regression, where we consider $Y_{o,d} \equiv$ I(N.Supp>0) and $Y_{o,d} \equiv$ I(N.Cust>0) as dependent variables. We present the standard OLS regression results in Table B.3 of Appendix B, in which we find similar but slightly smaller coefficients

[21] Please refer to Burchardi et al. (2019) for a more detailed discussion of the first-stage specification.

compared to those in the IV results, assuring that the IV strategy mitigates the potential downward bias of simple OLS regressions created by measurement errors or confounding factors.[22]

In Panel A of Table 2, we use I(N.Supp>0) as the dependent variable. In column (1), we regress I(N.Supp>0) on the log of ancestry in 2010 instrumented by the leave-out push-pull IVs from nine waves of US census data (1880-2000 waves) and control for origin and destination fixed effects. We find a significant and positive causal impact of immigration from origin country o to destination county d on global supply chain creation between firms located in d and supplier firms headquartered in o: A 10 percent increase in ancestry from origin o to destination d results in a 2.42 percentage point increase in the probability that at least one firm in destination d has a connection with origin o through supplier linkages.

From columns (2) to (7), we show the robustness of the result by adding various controls and more granular fixed effects. In column (2), we add the first five principal components of the higher-order interactions to the set of instruments. We find a slightly smaller magnitude but still a highly positive and significant impact of immigration on the global supply chain connection. In column (3), we add destination x continent-of-origin fixed effects ($\delta_d \times \delta_{c(o)}$) and origin x destination's census division fixed effects ($\delta_o \times \delta_{r(d)}$), exploiting only the variation within continents and within census divisions. Therefore, any common factors that operate within an origin-continent and destination-census-division pair are absorbed by these fixed effects.[23] The estimate in column (3) shows that a 10 percent increase in ancestry from origin country o to destination county d results in a 2.24 percentage point increase in the probability that at least one firm in destination d has a connection with origin o through supplier linkages. In other words, doubling the number of residents with ancestry from a given origin relative to the sample mean (from 320 to 640) increases by 4.9 percentage points the probability that at least one firm engages in a global supply chain relationship with a supplier company headquartered in that origin country.[24]

In column (4), we add an instrument constructed by using the 2000-2010 wave of migration. Column (5) adds the third-order polynomials of distance and latitude difference to capture any nonlinear relationship between global supply chain connection and distance, and column (6) adds measures of agricultural similarity between US counties and foreign countries to control for country-

[22] Burchardi et al. (2019) argue that the endogenous assignment of migrants within the US could create a downward bias of the OLS coefficient on FDI flow. A similar logic holds on the relationship between the endogenous assignment of migrants and global supply chain relationships, which could create a downward bias in the OLS estimates: Migrations could be driven by differences in factor endowments (i.e., differences in wages between o and d), while global supply chain linkages could be driven by similarities in factor endowments.

[23] For instance, migrants from Asia have tended to settle in certain areas, such as the Pacific census division. If we were to observe larger global supply chain relationships between Asian countries and US counties in the Pacific census division, then one may argue that any confounding pair-specific factors could drive the result. The destination x continent-of-origin fixed effects ($\delta_d \times \delta_{c(o)}$) and origin x destination's census division fixed effects ($\delta_o \times \delta_{r(d)}$) can alleviate this concern.

[24] With $\hat{\beta} = 0.224$ from column (3) in Table 2: $\text{I(N.Supp>0}|\text{Ancestry}_{o,d} = 640) - \text{I(N.Supp>0}|\text{Ancestry}_{o,d} = 320) = 0.224 \left[\ln\left(1 + \frac{640}{1000}\right) - \ln\left(1 + \frac{320}{1000}\right)\right] \approx 0.049$.

Table 2: Impact of Ancestry Composition on Global Supply Chain Linkages: IV Regression

Panel A	I(N.Supp>0)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Ancestry 2010	0.242***	0.194***	0.224***	0.218***	0.224***	0.233***	0.234***
	(0.019)	(0.023)	(0.020)	(0.020)	(0.020)	(0.023)	(0.020)
Log Distance	0.016	0.012	0.055	0.053	0.033	0.067	-0.002
	(0.015)	(0.014)	(0.037)	(0.037)	(0.044)	(0.058)	(0.051)
First-stage F stat	11.0	2448.0	162.2	195.4	158.1	102.8	186.2
Destination FE	✓	✓	-	-	-	-	-
Origin FE	✓	✓	-	-	-	-	-
Principal Components	-	✓	✓	✓	✓	✓	✓
Destination x Continent FE	-	-	✓	✓	✓	✓	✓
Origin x Census Division FE	-	-	✓	✓	✓	✓	-
$I^{2010}_{o,-r(d)}(I^{2010}_{-c(o),d}/I^{2010}_{-c(o)})$	-	-	-	✓	-	-	-
3rd order poly in dist and lat	-	-	-	-	✓	✓	-
Agricultural Similarity	-	-	-	-	-	✓	-
Origin x State FE	-	-	-	-	-	-	✓
Observations	612495	612495	612495	612495	612495	459150	612300
Panel B	I(N.Cust>0)						
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Log Ancestry 2010	0.232***	0.196***	0.207***	0.208***	0.208***	0.209***	0.213***
	(0.018)	(0.022)	(0.019)	(0.019)	(0.019)	(0.023)	(0.021)
Log Distance	0.017	0.014	0.041	0.041	0.029	0.050	-0.022
	(0.018)	(0.017)	(0.040)	(0.041)	(0.045)	(0.060)	(0.070)
First-stage F stat	11.0	2448.0	162.2	195.4	158.1	102.8	186.2
Destination FE	✓	✓	-	-	-	-	-
Origin FE	✓	✓	-	-	-	-	-
Principal Components	-	✓	✓	✓	✓	✓	✓
Destination x Continent FE	-	-	✓	✓	✓	✓	✓
Origin x Census Division FE	-	-	✓	✓	✓	✓	-
$I^{2010}_{o,-r(d)}(I^{2010}_{-c(o),d}/I^{2010}_{-c(o)})$	-	-	-	✓	-	-	-
3rd order poly in dist and lat	-	-	-	-	✓	✓	-
Agricultural Similarity	-	-	-	-	-	✓	-
Origin x State FE	-	-	-	-	-	-	✓
Observations	612495	612495	612495	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1). The main dependent variables are I(N.Supp>0) and I(N.Cust>0). I(N.Supp>0) is a dummy variable that equals one if any firm whose trade-engaging establishments located in destination county d has at least one supplier firm headquartered in origin country o between 2011 and 2014, and zero otherwise. I(N.Cust>0) is similarly defined, where we use customer firms headquartered in origin o instead of supplier firms. The key variable of interest is Log Ancestry 2010, which is instrumented using various specifications of Equation (3.2). All columns include $\{I^t_{o,-r(d)}(I^t_{-c(o),d}/I^t_{-c(o)})\}_{t=1880,\ldots,2000}$ as instruments. Columns (1)-(7) focus on I(N.Supp>0). Column (1) includes destination fixed effects and origin fixed effects. In column (2), we additionally include the first five principal components of the higher-order interactions of the push and pull factors as instruments. Column (3) includes "destination county"-by-continent fixed effects and "origin country"-by-"census division" fixed effects. Column (4) includes the interaction of the push and pull factor constructed using data from the 2006-2010 American Community Survey. Columns (5)-(6) add the 3rd-order polynomials in distance and latitude, and the measure of agricultural similarity. Column (7) includes "origin country"-by-"state" fixed effects instead of "origin country"-by-"census division" fixed effects. Columns (8)-(14) repeat the analyses using I(N.Cust>0). All regressions control for log distance and latitude difference. Standard errors are clustered at the origin-country level; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

county specific attractiveness driven by agricultural similarity, which could jointly affect migration and supply chain linkages between origin countries and destination counties. Finally, in column (7), we include more stringent fixed effects by including origin x destination's state fixed effects, exploiting only variation within US states. Across all specifications, we find a significant and positive causal impact of ancestry composition in 2010 on global supply chain connections through supplier-firm linkages for the period 2011-2014. We obtain similar first-stage results as those in Burchardi et al. (2019), confirming the relevance of the instruments: The Kleibergen-Papp first-stage F-statistics strongly reject the null of weak instruments across all specifications at the conventional level, where F-statistics range from 11 to 2448.

Panel B of Table 2 repeats the exercise by using I(N.Cust>0) as the dependent variable. Again, we find a significant and positive causal impact of ancestry composition on global supply chain connections through customer-firm linkages across all specifications. Table B.4 of Appendix B further shows the robustness of our results. Here, we consider a union of supplier-firm and customer-firm linkages, I(N.Link>0), which is formally defined as a dummy variable that equals one if any firm whose trade-engaging establishments are located in destination county d has *either* supplier firms *or* customer firms headquartered in origin country o. Based on column (7) of Table B.4, our estimate of interest shows that a 10 percent increase in ancestry from origin o to destination d results in a 2.03 percentage point increase in the probability that at least one firm in destination d has a connection with origin o through customer or supplier linkages. This finding implies that doubling the number of residents with ancestry from a given origin country relative to the sample mean (from 320 to 640) increases by 4.5 percentage points the probability that at least one firm engages in a global supply chain relationship with the customer- or supplier-firm headquartered in that origin country.

In Table B.5 of Appendix B, we explore the impact of ancestry on global supply chain linkages at the intensive margin. To do so, we consider new dependent variables—log of the number of supplier linkages (Log N.Supp) and customer linkages (Log N.Cust), respectively, conditional on having positive supply chain network connections between origin country o and destination county d. Across all specifications, we find a strong causal impact of ancestry composition on global supply chain formation even at the intensive margin. Based on column (3) (column (10)), the estimate implies that doubling the number of residents with ancestry from a given origin country relative to the sample mean (from 320 to 640) increases the number of supplier (customer) linkages by 5.5 percent (5.5 percent).

4.2 Robustness and Placebo Tests

In this section, we conduct further robustness checks to corroborate our findings. We first show that global supply chain formation is a distinct phenomenon from foreign direct investment (FDI): Directly controlling for FDI and strictly ruling out any potential within-multinational linkages do not alter our result. Second, we conduct a placebo test and show that the impact of ancestry composition on global supply chain linkages works only through trade-engaging establishments—which directly engage in firm-to-firm trade with foreign companies—and not through non-trade-engaging establishments. Third, we show that the results are robust to restricting firms to single-establishment firms or defining firms' US county locations based on their headquarters locations. Fourth, neither using ancestry compositions prior to 2010 (i.e., 1980, 1990, and 2000) nor restricting supply chain linkages to those newly formed after 2010 affects our main results, reassuring that reverse causality is not driving our core results. Our results are also robust to dropping Asian countries and US counties on the West Coast, implying that the results are not particularly driven by the strong tie between these regions. Also, we confirm that dropping top ancestry origins or top global supply chain partner countries do not alter our results. Finally, the results are robust to using a relative measure of ancestry normalized by country size.

4.2.1 Controlling for FDI and Excluding Potential Multinational Linkages

The IV estimates above will yield a biased estimate of the causal effect of ancestry if global supply chain relationships disproportionately occur between origin o and destination d pairs that also have more FDI links. Then, the estimates could capture the impact of ancestry on FDI rather than on global supply chain relationships. To check for this possibility, we additionally include an FDI dummy variable between origin country o and destination county d in Equation (3.1) and repeat the analysis.[25]

Table 3 presents the estimation results. We first find that FDI linkages between origin country o and destination county d, which are by nature intra-firm linkages through parents and subsidiaries, are positively associated with inter-firm linkages through global supply chain networks. The positive correlation between FDI and global supply chain relationships thus provides us with a rationale for controlling for FDI. Despite this positive correlation, even after controlling for the FDI dummy, we find a statistically significant and positive impact of ancestry composition on global supply chain linkages across all columns in Table 3. The magnitudes are somewhat smaller than the corresponding results in Table 2, potentially reflecting the confounding effects between FDI and global supply chain relationships. Overall, we fail to reject the null impact of ancestry on global supply chain relationships.

Based on column (1) (column (5)), doubling the number of residents with ancestry from a given origin relative to the sample mean (from 320 to 640) increases, by 3.9 percentage points (3.5 percentage points) the probability that at least one firm engages in a global supply chain relationship

[25]We use the FDI dummy provided by Burchardi et al. (2019), which takes a value of one if at least one firm in destination d has at least one parent or subsidiary in origin o in 2007. We thank the authors who have publicly shared their datasets on their webpages and the journal's data archives.

Table 3: Controlling for FDI

	I(N.Supp>0)				I(N.Cust>0)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Ancestry 2010	0.179***	0.179***	0.192***	0.190***	0.162***	0.162***	0.168***	0.169***
	(0.019)	(0.019)	(0.022)	(0.020)	(0.019)	(0.019)	(0.022)	(0.020)
FDI Dummy	0.237***	0.237***	0.219***	0.221***	0.237***	0.237***	0.226***	0.224***
	(0.023)	(0.023)	(0.023)	(0.022)	(0.022)	(0.022)	(0.023)	(0.021)
First-stage F stat	76.3	75.9	60.2	84.8	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	612495	612495	459150	612300	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we additionally control for the FDI dummy from Burchardi et al. (2019). The FDI dummy takes a value of one if at least one firm in destination county d has at least one parent or subsidiary in origin country o in 2007. Columns (1)-(4) use the specifications in columns (4)-(7) in Table 2, while columns (5)-(8) use the specifications in columns (11)-(14) in Table 2.

with a supplier (customer) company headquartered in that origin country.[26]

In Table B.6 of Appendix B, we further show the robustness of our results by strictly restricting US firms to those firms with headquarters located in the US. Therefore, in this exercise, we restrict supplier-customer relationships to those between foreign firms whose headquarters are located *outside the US* and US firms whose headquarters are located *in the US*.[27] Since the foreign suppliers or customers of these firms have headquarters outside the US, we strictly rule out any potential supplier-customer linkages that might capture within-multinational relationships between headquarters and their foreign subsidiaries.

4.2.2 Placebo Test: Non-trade-engaging Establishments

In our baseline analyses, to measure the location of firms in US counties, we restrict establishments to those that engage in international trade between 2011 and 2014. Given that our firm-to-firm

[26] Specifically, with $\hat{\beta} = 0.179$ from column (1) in Table 3: $\text{I(N.Supp>0}|\text{Ancestry}_{o,d} = 640) - \text{I(N.Supp>0}|\text{Ancestry}_{o,d} = 320) = 0.179 \left[\ln\left(1+\frac{640}{1000}\right) - \ln\left(1+\frac{320}{1000}\right)\right] \approx 0.039$.

[27] In the baseline analysis, we define US firms as companies that have establishments located in the US, regardless of their headquarters location. Therefore, it is possible that these firms have headquarters outside the US. In the data, among all firms that have at least one trade-engaging establishment in the US, only 16% have headquarters outside the US. Table B.6 shows that our results are robust to strictly restricting US firms to those who have headquarters in the US.

Table 4: Placebo Test: Establishments Not Engaging in Trade between 2011 and 2014

	I(N.Supp>0)				I(N.Cust>0)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Ancestry 2010	0.006	0.006	0.014	0.003	0.030	0.030	0.013	0.022
	(0.020)	(0.020)	(0.025)	(0.020)	(0.021)	(0.021)	(0.024)	(0.022)
First-stage F stat	76.3	75.9	60.2	84.8	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	612495	612495	459150	612300	612495	612495	459150	612300

Notes. This table presents the regression results of the placebo test. Specifically, we present the coefficient estimates from the IV regressions of Equation (3.1), where we measure firms' US county location solely based on non-trade-engaging establishments. Columns (1)-(4) use the specifications in columns (1)-(4) in Table 3, while columns (5)-(8) use the specifications in columns (5)-(8) in Table 3.

global supply chain information is provided at the firm level—and not at the establishment level—we presume that a firm's establishments that engaged in international trade between 2011 and 2014 (trade-engaging establishments) are more likely to be connected with foreign firms through global supply chain relationships between in that period, compared to the same firm's establishments that did not engage in trade in that period (non-trade-engaging establishments).[28] For example, if a firm that has a supplier firm in Italy has establishments in Sacramento county (California) and San Augustine county (Texas), and if the establishment in Sacramento engaged in international trade while that in San Augustine did not, we would expect that the establishment in Sacramento is more likely to be connected with the firm's supplier in Italy.

This provides us with a natural placebo test, where we measure firms' US county locations solely based on non-trade-engaging establishments. Table 4 presents the estimation results. Columns (1)-(4) show that there is no impact of ancestry composition on supplier linkages between origin country o and destination county d if we measure firms' US county location using non-trade-engaging establishments. This result means that, for example, even if a firm has an establishment in San Augustine, if that establishment does not engage in trade, then an (exogenous) increase in Italian ancestry in San Augustine does not increase global supply chain linkages between that county (d) and Italy (o) through supplier linkages.

[28] Such a complication does not exist if a firm has a single establishment. We show the robustness of our results (i) by restricting firms to single-establishment firms and (ii) by measuring the location of firms using the county location of headquarters establishments. See Table B.7 of Appendix B.

Columns (5)-(8) repeat the exercise using customer linkages. Although columns (5) and (6) give weakly positive coefficients, the effects are economically small. Additionally, controlling for agricultural similarity (column (7)) and more granular fixed effects (column (8)) makes the estimates statistically insignificant, confirming that the impact of ancestry composition on firm-to-firm supply chain linkages works primarily through trade-engaging establishments in the US and their connection with foreign customer- or supplier-firms.

4.2.3 Alternative Ways of Defining Firms' US County Locations

Given that our firm-to-firm global supply chain information is provided at the firm level, while location information for US firms is provided at the establishment level, multi-establishment firms create complications regarding how we should assign firm-to-firm linkages to each establishment. Our baseline analyses used trade-engaging establishments to measure firms' US county locations.

As a robustness check, we investigate whether our results are robust to (i) restricting firms to single-establishment firms and (ii) measuring the location of firms using the county location of their headquarters establishments. As Table B.7 in Appendix B shows, we find robust results on these specifications.

4.2.4 Reverse Causality

We also directly address the potential reverse causality concern in two ways. First, we use ancestry compositions in 2000, 1990, and 1980, instead of using that in 2010. Second, we strictly restrict global supply chain linkages in 2011-2014 to those that did not exist in 2010 but were newly formed in 2011-2014. Together, the robustness of our results under these alternative specifications reassures that the core results are not driven by reverse causality.

Using Ancestry Compositions Prior to 2010 We also show that our results are robust to using ancestry compositions prior to 2010. Global supply chain linkages between 2011 and 2014 might have existed for a prolonged time period, potentially prior to 2010. Our main specification in Equation (3.1), therefore, uses IVs that exploit historical migration spanning from 1880 to 2000 (except for columns (4) and (11), which use the 2006-2010 wave). We additionally show that the ancestry composition that dates back to 1980 can explain the current global supply chain linkages, reassuring the role of historical migration in the formation of global supply chain connections.

Specifically, in Table B.8 of Appendix B, we consider the log of ancestry in 2000, 1990, and 1980, where we exclude IVs that are constructed by using the 1990 and 2000 waves (as well as principal components). We find robust results under these alternative specifications.[29]

[29] We tend to obtain slightly low F statistics (approximately 11) mainly because we do not use the principal components.

Restricting Global Supply Chains to Newly Formed Linkages after 2010 We further show the robustness of our results by restricting global supply chain connections to those newly formed after 2010. Specifically, we measure linkages based on global supply chain connections that did not exist in 2010 but were newly formed between 2011 and 2014. Table B.9 presents the results and confirms the robustness of our findings.

4.2.5 Excluding West Coast Counties and Asian Countries

We show that the results are robust to dropping countries in Asia and US counties on the West Coast, implying that the results are not particularly driven by the strong tie between these regions. For example, tech companies in Silicon Valley have been outsourcing their production to countries in Asia and have also been sourcing key intermediate goods (e.g., semiconductors) from such countries. Table B.10 shows that our main findings hold after dropping countries in Asia and US counties in the West census region.

4.2.6 Excluding Top Ancestry Origins and Global Supply Chain Partner Countries

We further investigate whether our results are particularly driven by a subset of countries, namely top ancestry origins or top global supply chain partner countries. In Table B.11, we drop top five ancestry origins—Germany, United Kingdom, Mexico, Ireland, and Italy. Columns (1) to (5) drop each country one by one, and Column (6) drops all five countries. We find statistically significant estimates across all columns. The estimated coefficients remain remarkably stable across columns, with the exception of dropping Mexico. Dropping Mexico results in slightly smaller—yet highly significant—estimated coefficients of 0.149 (for supplier linkage) and 0.136 (for customer linkage) compared to their counterparts in Table 3—0.179 (for supplier linkage) and 0.162 (for customer linkage), respectively. This reveals the importance of Mexico at the origin-destination level. Although Mexico is the third ancestry origin in terms of the total number of ancestry (Appendix Table B.1), nine among top 10 origin-destination-level ancestry consist of Mexico—mostly paired with counties in California and Texas. Despite the importance of Mexico, the robust results in Table B.11 show that our main finding is not particularly driven by a few important countries.

In Table B.12 and Table B.13, we repeat the analyses using (i) Canada plus top five Asian countries—China, Philippines, India, South Korea, and Japan—in terms of ancestry origins and (ii) top five global supply chain partners—United Kingdom, Japan, Canada, Germany, and France—as further robustness checks. All results indicate that our findings are not particularly driven by a subset of important countries in terms of ancestry origin or global supply chain partnership.[30]

[30] In unpublished analyses, we also confirmed similar robust results at the intensive margin.

4.2.7 Alternative Measure of Ancestry Composition

Finally, we show the robustness of the results by using an alternative measure of ancestry composition. Note that it is more likely for a bigger country to have larger number of firms and that such a big country could serve as a major immigrant sourcing country to the US (due to its large population). Our baseline analyses handle such a concern by using granular fixed effects—i.e., origin fixed effects and destination fixed effects, or in a more demanding specification, origin-by-destination's census division fixed effects and destination-by-continent-of-origin fixed effects. For example, origin fixed effects allow us to effectively compare the impact of immigration from the same origin country to different US counties, thereby controlling for any country-specific effects such as country size.

Nevertheless, as a robustness check, we consider a *relative* measure of ancestry composition that directly takes into account country size differences. Specifically, we define the ancestry composition as the ratio of the number of ancestry to country population in 2010. Table B.14 shows that our main findings are robust under this alternative measure.

4.3 Beyond Conventional Perspectives on Supply Chain Relationships

In this section, we present two novel results that go beyond conventional perspectives on supply chain network relationships. First, we show that the impact of ancestry composition extends to strategic partners, which is a broader type of relationship that is related to but does not coincide with conventional supplier-customer relationships in global supply chains. Second, we show that the impact of ancestry composition on global supply chain linkages operates not only through manufacturing industries but also through non-manufacturing industries and that the impact through non-manufacturing industries is stronger.

4.3.1 Impact on Strategic Partners

Our baseline analyses focus on conventional global supply chain linkages: firm-to-firm networks that capture the supply of goods and services from one firm to another. FactSet Revere, however, provides relationship information that goes beyond conventional global supply chain networks: strategic partner relationships.[31] Strategic partnerships—joint ventures, research collaborations, marketing,

[31] FactSet Revere also provides the disclosed competitors of a firm, through which we can identify competitor relationships. We present results for competitor relationships in Table B.15 of Appendix B. Although it may be plausible that an increase in ancestry in a given county creates competition among firms—by making competitor firms in origin country o join markets in destination county d—we do not find strong evidence for this. The estimates are positive—consistent with such a view—but only marginally significant. However, this does not necessarily mean that an increase in ancestry does not place any competitive pressure on firms; instead, it may reflect larger measurement errors in competitor information relative to other types of relationship information. First, the competitor relationships reported by FactSet Revere are relatively small compared to other types of relationships: More than 50% are supplier-customer relationships, 30% are strategic partner relationships, and less than 20% are competitor relationships. Second, in contrast to supplier-customer relationships, where SEC 10-K filings—one of the important sources from which FactSet Revere collects supply chain relationship information—mandate publicly listed firms to disclose their major

Table 5: Impact of Ancestry Composition on Strategic Partner Relationships

	I(N.Link>0)			
	(1)	(2)	(3)	(4)
Log Ancestry 2010	0.040**	0.040**	0.041**	0.043**
	(0.018)	(0.018)	(0.019)	(0.019)
First-stage F stat	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-
Agricultural Similarity	-	-	✓	-
Origin x State FE	-	-	-	✓
Observations	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we use a linkage dummy based on strategic partner relationships as a dependent variable. Columns (1)-(4) use the specifications in columns (1)-(4) in Table 3, where only the dependent variable is replaced with the linkage dummy based on strategic partner relationships.

integrated product offerings, in- and out-licensing, etc.—could be viewed as a broader type of firm-to-firm relationships (i.e., partnerships) that partially overlap with but do not exactly fall into supplier-customer relationships.[32]

Columns (1)-(4) in Table 5 reveal that there is a statistically significant and positive causal impact of ancestry composition on strategic partnership between origin country o and destination county d.[33] Thus, our main results extend to a broader type of relationship, beyond conventional global supply chains.

4.3.2 Manufacturing vs. Non-manufacturing Industries

In our main analysis, we do not distinguish between (trade-engaging) manufacturing and non-manufacturing establishments. We explore whether our results are driven by (trade-engaging) manufacturing establishments or non-manufacturing establishments or both. To this end, we perform separate analyses by restricting establishments to manufacturing (SIC 2-digit 20-39) and

customers—those who account for more than 10% of their revenues—information on competitor relationships is collected purely based on voluntary disclosure through public sources, investor presentations, and press releases. Therefore, some competitor relationships may not be captured by the dataset. This could create a downward bias of estimates due to measurement errors.

[32]We follow the classification provided by FactSet Revere and define two firms as having a strategic partnership if one of them reports that it has the following type of relationship with the other: joint venture, research collaboration, marketing, integrated product offering, in- and out-licensing, equity investment, and product manufacturing and distribution.

[33]As in Table 2, we use trade-engaging establishments to identify the county locations of each firm in the US.

non-manufacturing.

Despite the fact that the literature has focused primarily on goods trade in manufacturing industries (potentially due to limited accessibility to service trade data), we identify the impact of ancestry composition on global supply chain formation for both manufacturing and non-manufacturing industries. In fact, the impact is greater for non-manufacturing industries.

Panel A in Table 6 measures firms' US county location solely based on the location of trade-engaging *manufacturing* establishments. That is, the dependent variable I(N.Supp>0) (similarly, I(N.Cust>0)) is equal to one if any firm that has at least one trade-engaging *manufacturing* establishment located in destination county d has at least one supplier (customer) firm headquartered in origin country o. Similarly, Panel B measures firms' US county location using the location of trade-engaging *non-manufacturing* establishments. We find that the impact of ancestry on global supply chain relationships works through both manufacturing and non-manufacturing establishments and that the impact is stronger for non-manufacturing establishments.

In Appendix Table B.16, we show that the results are similar if we further restrict non-manufacturing establishments to those strictly operating in service-related sectors.[34] This finding implies that the impact of ancestry composition on global supply chain formations operates not only in goods trade but also in service trade.

5 Exploration of Mechanisms

Thus far, we have found robust evidence that the presence of co-ethnic networks that have been shaped by immigrants to the US for more than a century can causally explain the current structure of global supply chain relationships. A natural follow-up question is then *why* firm-to-firm networks are more prevalent between a US county and a foreign country when co-ethnic networks are more predominant between them. In this section, we now delve into the specific mechanisms responsible for the linkage between global supply chains and co-ethnic networks.

5.1 Barriers to International Firm-to-Firm Relationships

It is well documented in the literature that international transactions, including global supply chain relationships, are riskier than domestic transactions. For instance, international transactions usually involve greater geographical distances than do domestic transactions, which require more working capital for firms engaged in international transactions due to longer transportation times. Moreover, language barriers and information and contractual frictions are more severe compared to those in domestic transactions. Thus, firms need to put extra effort into understanding the differences in

[34] Specifically, we measure firms' US county location using trade-engaging non-manufacturing establishments classified as (i) transportation & public utilities (SIC 40-49); (ii) finance, insurance, and real estate (SIC 60-67); and (iii) services (SIC 70-89).

Table 6: (i) Manufacturing Establishments, (ii) Non-manufacturing Establishments

Panel A	Manufacturing Establishments							
	I(N.Supp>0)				I(N.Cust>0)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Ancestry 2010	0.161***	0.161***	0.176***	0.174***	0.146***	0.147***	0.159***	0.154***
	(0.020)	(0.020)	(0.022)	(0.021)	(0.022)	(0.022)	(0.025)	(0.025)
First-stage F stat	76.3	75.9	60.2	84.8	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	612495	612495	459150	612300	612495	612495	459150	612300

Panel B	Non-manufacturing Establishments							
	I(N.Supp>0)				I(N.Cust>0)			
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Log Ancestry 2010	0.204***	0.204***	0.218***	0.216***	0.198***	0.198***	0.206***	0.207***
	(0.021)	(0.021)	(0.026)	(0.022)	(0.021)	(0.021)	(0.025)	(0.022)
First-stage F stat	76.3	75.9	60.2	84.8	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	612495	612495	459150	612300	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where Panel A measures firms' US county location using trade-engaging manufacturing establishments and Panel B measures this location using trade-engaging non-manufacturing establishments. Columns (1)-(4) in Panel A and columns (9)-(12) in Panel B use the specifications in columns (1)-(4) in Table 3, while columns (5)-(8) in Panel A and columns (13)-(16) in Panel B use the specifications in columns (5)-(8) in Table 3.

institutional contexts across countries in international transactions. In these frictional environments, Rauch (2001) noted that co-ethnic networks that operate across national borders may help mitigate the above-mentioned kinds of barriers.

Among many types of international transactions, a firm-to-firm supply chain relationship has unique features that distinguish it from within-firm transactions such as FDI—the subject of the study in Burchardi et al. (2019). While the global supply chain network is characterized by an arm's-length firm-to-firm relationship, FDI is conducted within the boundaries of the firm. In other words, the boundaries of a firm are defined in a different manner (Coase, 1937; Grossman and Hart, 1986; Hart and Moore, 1990), and the friction of incomplete contracts tends to be more severe in an arm's-length relationship than in integrated companies (Williamson, 1975, 1985; Antras and Helpman, 2004).

Given the risk surrounding international transactions, financial or credit constraints may critically hamper firms' participation in global supply chain networks (Kim and Shin, 2012; Carluccio and Fally, 2012; Basco, 2013; Choi, 2020). For example, due to arm's length transactions in global supply chains, international firm-to-firm relationships rely heavily on informal finance, such as trade credit, to address the mismatch between payment and delivery in international transactions.[35] The mismatch can arise because a supplier (i) can require the customer to pay for goods before they are loaded for shipment or (ii) can allow the customer to pay at some moment in time after the goods have arrived at their destination (Antras and Foley, 2015). While domestic transactions also involve financial claims between firms through trade credit, cross-border financial claims are associated with more risks, owing to their longer transportation times and different institutional contexts across countries.[36] More generally, financial frictions and credit constraints may negatively affect general business operations, which could dampen active international transactions.

5.2 Co-ethnic Networks, Trust, and Credit Constraints in Global Supply Chains

One way to overcome credit frictions is to rely on informal finance, an important feature of which is that social trust plays a major role in granting and receiving credit,[37] especially when there is no formal contractual enforcement mechanism in place (Fafchamps, 2000; Karlan, 2005; Wu et al., 2014; Levine et al., 2018). Due to the importance of social trust in facilitating informal finance, co-ethnic networks often play an important role.

This situation is well illustrated by the history of Korean immigrants who have settled in Los Angeles, the US city with the largest Korean-American population. A survey of Korean immigrants in Los Angeles by Min (1993) shows that Korean entrepreneurs who immigrated and settled in Los Angeles mostly engaged in trading activities, such as trade in fashion items and garments, with vendors in South Korea. Through the advantages associated with their language and ethnic background, many Korean immigrants have been able to establish import businesses dealing in

[35]Trade credit is a financial instrument such that credits are extended bilaterally between non-financial firms in transactions. Ahn (2021) reviewed the related literature and summarized the major payment methods in international transactions, among which trade credit accounts for 65 to 96% of international transactions.

[36]Note that, while cross-border financial claims typically arise in firm-to-firm relationships, FDI is conducted within the same entity. Therefore, multinationals may help to relax financial frictions in international transactions (Antras et al., 2009; Desai et al., 2004), and due to such an advantage, multinational affiliates may fare better at export performance in financially vulnerable environments (Manova et al., 2015).

[37]Fukuyama (1996) described social trust as *"the expectation that arises within a community of regular, honest, cooperative behavior, based on commonly shared norms, on the part of other members of that community."*

Korean-imported goods (e.g., textiles, wigs), in which Korean importers processed and distributed Korean-made products mainly to Korean wholesalers, who in turn distributed to other Korean retailers.

A notable aspect behind such development is that Korean immigrants have been able to form trust-based rotating credit associations (RCAs)—known as "kye"—to overcome credit constraints (Light et al., 1990).[38] It has been documented that RCAs facilitate the entrepreneurship of immigrants and ethnic minorities who settle in developed countries and have been widespread in the majority of US immigrant communities that have originated from emerging or developing countries (Light and Pham, 1998). As immigrants or ethnic minorities lack credit ratings or collateral, it is difficult for them to access the conventional financial system in the US. By relying on RCAs or an ancient Asian lending practice exemplified by "kye", Korean immigrants were able to overcome such financial obstacles and finance their businesses, which even enabled them to engage in trading activities with Korean exporters. According to the *Los Angeles Times* article,[39] kye was widely used among approximately 400,000 Korean immigrants in the late 1980s, which supported the rapid expansion of the LA apparel industry in Koreatown in this period.[40]

As such, social trust is an essential element in informal financing relationships, which resolves the mismatch between payment and delivery. Because this mismatch gives rise to an implicit financing contract such that one party promises to pay at a later date, it naturally involves default risk. In informal finance relations, as there are typically no collaterals and/or guarantees from formal financial institutions, trust plays an important role in mitigating default risks.[41] Furthermore, unlike in domestic relationships, in global supply chain relationships, social trust can be even more important when a firm interacts with other firms in countries with poorly functioning institutions (e.g., poor contract enforcement).

Karlan et al. (2009) argue that network connections between individuals can be used as social collateral (or network-based trust) to secure informal borrowing. We argue that one such social

[38] RCAs—or, more generally, rotating savings and credit associations (ROSCAs)—refer to informal social networks whose participants agree to make financial contributions to a fund periodically (Besley et al., 1993; Anderson and Baland, 2002). A good example of an RCA is a so-called "kye"—in which a group of a dozen or more friends or associates gathers monthly, and each member contributes the same amount, usually ranging from \$100 to \$50,000, to a common pot. Each month, a different member takes the "kitty" and agrees to pay interest to the others. The members also promise to remain in the kye until each has collected the pot.

[39] *Los Angeles Times* (by D. Frantz, Oct. 1988), Hanmi Bank Uses Ancient Asian Lending Practice to Help Koreans (link to the article).

[40] Since the early 1970s, Korean immigrants have played a major role in the apparel industry in Los Angeles, leveraging their co-ethnic ties back home to engage in textile trades as South Korea industrializes largely through the garment sector and its export to the US during the 1960s and 1970s. Over the years, the industry has expanded, and approximately 3,000 businesses are officially registered with the Korean Apparel Manufacturers Association of Los Angeles. *Pacific Standard* (by Christina Moon, Mar. 2017), The Secret World of Fast Fashion (link to the article).

[41] In a related manner, as Guiso et al. (2004) noted, *"Since financial contracts are the ultimate trust-intensive contracts, social capital should have major effects on the development of financial markets. Financing is nothing but an exchange of a sum of money today for a promise to return more money in the future. Whether such an exchange can take place depends not only on the legal enforceability of contracts but also on the extent to which the financier trusts the financee."*

network that enables each involved party to build trust is a co-ethnic network. One reason that co-ethnic networks facilitate trust might be that they communicate more effectively within co-ethnic societies. For instance, Ederer and Schneider (2022) show that communication raises trust in their laboratory and online trust experiment. Therefore, co-ethnic networks may facilitate trade credit provision by enhancing trust in trade credit relationships (Fafchamps, 2000).[42]

To summarize, credit frictions could hamper those international firm-to-firm transactions that rely heavily on trade credit—which is riskier than a domestic transaction due to longer times and the potential absence of formal contract enforcement procedures—and more generally, by negatively affecting business operations. Social networks (such as co-ethnic networks) can enhance trust, which serves as social collateral, to relax financial frictions and grease the wheels of granting and receiving trade credit in global supply chain relationships.

5.2.1 Testable Hypothesis

Therefore, we would expect to observe that the positive impact is stronger for US counties in which more credit-constrained firms are located. If so, the presence of co-ethnic networks can serve as social collateral to overcome the credit constraints of customer and supplier firms in global supply chain relationships. To test our hypothesis, we set up the following empirical specification:

$$Y_{o,d} = \delta_o + \delta_d + \beta A_{o,d}^{2010} + \gamma A_{o,d}^{2010} \times CC_d + X'_{o,d}\eta + \varepsilon_{o,d} \tag{5.1}$$

where CC_d denotes destination-level credit constraints, which measures the average credit constraint of establishments located in destination county d.[43] A higher value of CC, i.e., credit constraint, indicates *worse* trade credit solvency. To be more specific, CC is measured by (100-PayDex) so that a higher CC indicates lower PayDex scores, and thus, higher credit constraints. For the measures of PayDex, we use PayDexMax and PayDexMin, which stand for yearly maximum and minimum PayDex scores, respectively.[44] To facilitate the interpretation of the coefficients, we standardize CC so that the sample mean equals zero and the sample standard deviation equals one.

[42] In various contexts, the literature documents that co-ethnicity matters for the formation of business relationships such as those between founders and venture capitalists (Bengtsson and Hsu, 2015), between inventors and the foreign country with which the inventors' firm builds a multinational affiliate or engages in business cooperation (Foley and Kerr, 2013; Xie et al., 2022), and between Indian hirers and workers on an online outsourcing platform (Ghani et al., 2014).

[43] CC_d is obtained by calculating a weighted average of establishment-level credit constraint measures, averaged across establishments within each destination. We use establishments' employment as the weight of each establishment.

[44] PayDex score from NETS, which is originally derived from D&B, is particularly well-suited for testing our hypothesis since it is directly based on actual trade credit performance at the establishment level. Suppliers of goods and services around the world report to D&B their trade credit experiences with each business, and D&B—from which NETS data are derived—constructs the PayDex score based on this information. Since D&B is one of the largest credit rating companies in the world, it has a strong incentive to construct accurate credit ratings, and suppliers around the world use the PayDex score of their potential trading partners to decide whether to extend trade credit to them. For this reason, the PayDex information in NETS data has been fruitfully used in the context in which trade credit solvency may be important (e.g., Akey and Appel, 2021; Borisov et al., 2021).

Table 7: Heterogeneous Treatment Effect: Credit Constraints

	100-PayDexMax		100-PayDexMin	
	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)	(3)	(4)
Log Ancestry 2010	0.035	0.064*	0.045	0.077**
	(0.035)	(0.035)	(0.039)	(0.033)
Log Ancestry 2010 × CC	0.216***	0.159***	0.190***	0.128***
	(0.053)	(0.052)	(0.055)	(0.042)
First-stage F stat	606.4	606.4	525.2	525.2
Destination FE	✓	✓	✓	✓
Origin FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	592995	592995	592995	592995

Notes. This table presents the heterogeneous treatment effect results by including the interaction of Log Ancestry 2010 with credit constraint measures. The regression equation is given by Equation (5.1). Columns (1) and (3) use the specifications in Panel A in Table 2, and columns (2) and (4) use the specifications in Panel B in Table 2, where we add the interaction between Log Ancestry 2010 and CC. CC, which is defined at the destination level, measures the average credit constraint of the establishments within each destination county. Columns (1)-(2) define CC using 100-PayDexMax; columns (3)-(4) define CC using 100-PayDexMin. All regressions control for the FDI dummy as in Table 3. To facilitate the interpretation of coefficients, all credit constraint variables are standardized so that the sample mean equals zero and the sample standard deviation equals one.

5.2.2 Validating the Hypothesis: Co-ethnic Networks Serve as Social Collateral

Table 7 confirms the hypothesis. The interaction term in the second row is the main focus of the empirical analysis. All coefficients are positive and statistically significant, meaning that when credit constraints are more serious, the positive impact of ancestry on foreign supply chain relationships becomes stronger. This result is consistent with our hypothesis that co-ancestral connections help mitigate credit constraints in global supply chain relationships.

In Table B.17 of Appendix B, we use the external financial dependence indicator by Rajan and Zingales (1998) as an alternative variable to capture destination-level credit constraints (i.e., CC_d).[45] The coefficients of the interaction term in the second row of Table B.17 are all positive and statistically significant, which means that when a county's industries are more dependent on external finance, the positive impact of co-ethnic networks on global supply chain networks becomes more pronounced, which corroborates the main result in Table 7.

We employ additional robustness checks to strengthen the validation of the hypothesis. In

[45] This approach is methodologically akin to that of Manova et al. (2015), who interact this measure with a multinational status dummy to find that being a multinational affiliate alleviates the negative impact of financial frictions on export performance. Similarly, we interact this measure with the ancestry composition to test if having co-ethnic networks helps dampen the negative impact of financial frictions on global firm-to-firm supply chains.

Appendix Tables B.18 and B.19, we further show that the results hold after controlling for other interaction terms, such as the interaction of log ancestry with origin-destination-level log distance, origin country-level ethnic fractionalization, destination county-level average firm size and the share of the final goods sector.[46,47] Consistent with Burchardi et al. (2019), we find a role of immigration on relaxing other trade barriers such as trade costs or information frictions. Longer distance may imply higher trade costs and higher information transmission costs. Information costs may be higher also for countries that are ethnically diverse and fractionalized. We find that the role of immigration on global supply chain formation becomes more important as distance between origin and destination increases and as ethnic fractionalization in origin country increases. However, Table B.18 reveals that the credit constraint channel remains valid even after controlling for this mechanism.

Furthermore, in Appendix Tables B.20–B.22, we explicitly allow for a firm dimension. That is, we perform analysis at the firm-origin-destination level and study the heterogeneous treatment effects.[48] The results are still robust under this firm-level specification: (i) A firm in a given county increases its foreign supply chain relationships with a given country if there is an exogenous increase in ancestry between these two regions, and (ii) when a *firm's* credit constraints are more serious in a given county, the positive impact of co-ethnic networks on the *firm's* foreign supply chain relationships becomes stronger; such heterogeneous effects are robust after controlling for other interaction terms.

Finally, given the importance of the interaction of co-ethnic networks and credit constraints in destination counties, a natural question that may arise is whether co-ethnic networks interact with financial frictions *in origin countries* and facilitate global supply chain formation. For example, Manova et al. (2015) show that multinationals in China are more likely to operate in more financially

[46]Ethnic fractionalization is defined as 1 minus the Herfindahl index of ethnicities in the origin country calculated using the data in Alesina et al. (2003). The share of the final goods sector is defined as follows. We first assign each establishment with the upstreamness index provided by Antras et al. (2012). Then, following Burchardi et al. (2019), we label an establishment as operating in the final goods sector if its upstreamness is less than 2. We then calculate the employment share of the final goods sector establishments for each county.

[47]The results demonstrate that the credit constraint channel operates even after controlling for other potential channels such as the role of trade costs or information frictions (proxied by distance and origin country's ethnic fractionalization), as well as the role of firm size and taste similarities between origin and destination (proxied by final goods sector share).

[48]Specifically, we run the following regression:

$$Y_{f,o,d} = \delta_f + \delta_{o,s} + \delta_{d,s} + \beta A_{o,d}^{2010} + \gamma A_{o,d}^{2010} \times CC_{f,d} + X_{o,d}'\eta + \varepsilon_{f,o,d}$$

$Y_{f,o,d}$ can take I(N.Supp>0), I(N.Cust>0), and I(N.Link>0). I(N.Supp>0) and I(N.Cust>0) are dummy variables that equal one if firm f, who has a trade-engaging establishment in destination county d, has at least one supplier (customer) headquartered in origin country o. I(N.Link>0) is a linkage dummy that incorporates both supplier and customer linkages, as in Table B.4. $CC_{f,d}$ measures firm-destination-level credit constraints, which is obtained by calculating the weighted average of "100-PayDex" across trade-engaging establishments within each firm-destination, weighted by establishments' employment. δ_f indicates firm fixed effects. $\delta_{o,s}$ ($\delta_{d,s}$) indicates origin (destination)-by-sector fixed effects, where sector is defined as a firm's primary SIC 4-digit industry. These fixed effects absorb any firm-specific or origin (destination)-by-sector-specific (un)observed components. The goal of the exercise is to check if higher ancestry composition from origin country o to destination county d increases the probability that firm f—who has a trade-engaging establishment in destination d—has supply chain linkages with origin o between 2011-2014.

vulnerable sectors, in which they exhibit better performance than local companies. This implies that the foreign owned firms in China do not necessarily need to rely on local financial systems compared to the local firms. Similarly, co-ethnic networks could facilitate international firm-to-firm transactions by resolving financial vulnerability issues in origin countries. On the other hand, if a country has bad institutions and weak judicial quality, it is also possible that immigration from that particular country suppresses formation of international transactions rather than facilitating them. For example, local businesses in destination counties that experienced a large inflow of immigrants from a country with bad institutions may be reluctant to open a business relationship with the origin country as information on the danger of doing business with firms therein may spread out in the local society. Given that countries with better judicial quality and legal development is likely to be positively associated with financial development, co-ethnic networks could dampen international transactions if foreign firms are operating in countries with weak financial development.

We test this hypothesis in Appendix Tables B.23 and B.24 by using a country-level measure of financial development by Beck (2002) and Manova (2013).[49] First, Table B.23 shows that there is no significant interaction between log ancestry and origin country's *financial weakness*—which is defined as the "minus" financial development (-FD)—potentially reflecting the two aforementioned opposing forces. However, as Table B.24 reveals, once we additionally control for the interaction between log ancestry and origin country's *judicial weakness*, which is defined as the "minus" judicial quality (-JQ)[50], we find a statistically significant positive interaction of log ancestry and origin country's financial weakness. This implies that, conditional on the interactions of common ancestry and judicial quality in origin countries, co-ethnic networks alleviate financial weakness of origin countries and facilitate global supply chain formation, just as the way co-ethnic networks relax credit constraints in destination counties.[51]

5.2.3 Solidifying the Mechanism: Role of Contractual Enforcement

Given that the positive impact of immigration on global supply chain formation is stronger in destination counties with higher credit constraints, does such an interaction—between co-ethnic networks and credit constraints—further interact with origin country characteristics in which foreign

[49]Specifically, the measure of financial development is defined as the amount of credit by banks and other financial intermediaries to the private sector as a share of GDP.

[50]Country-level measure of judicial quality comes from Nunn (2007). We revisit this measure in more depth in Section 5.2.3.

[51]Appendix Table B.25 shows the result where we only include interactions of log ancestry with credit constraints (in destination) and judicial weakness. Similar to Table B.24, we broadly find significant negative coefficients for the interaction of log ancestry and judicial weakness. This implies that common ancestry and good institutions are complements rather than substitutes, which is also consistent with the findings by Burchardi et al. (2019). Although the impact of immigration on global supply chain formation becomes weaker under bad institutions (i.e., higher (-JQ)), we show that the role of immigration on alleviating credit constraints in destination counties becomes more stronger if origin countries have weaker judicial quality (i.e., higher (-JQ)) by using triple interaction specification in the next section (Section 5.2.3).

firms operate? One prominent factor that could interplay with co-ethnic networks and credit constraints is the judicial quality of the origin country. For example, Antras and Foley (2015) show that the formation of trade relationships between exporters and importers mitigates credit constraints in international transactions, especially for countries with weak contract enforcement. Similarly, in global supply chain relationships, we expect to observe that the positive impact of immigration on global supply chain formation is not only stronger for credit-constrained firms in US counties, but that such a positive interaction becomes more amplified if US firms interact with firms in foreign countries with weak contract enforcement.[52] To test this additional hypothesis, we set up the following empirical specification:

$$Y_{o,d} = \delta_o + \delta_d + \beta_1 A_{o,d}^{2010} + \beta_2 A_{o,d}^{2010} \times CC_d + \beta_3 A_{o,d}^{2010} \times CC_d \times (-JQ_o) + \beta_4 A_{o,d}^{2010} \times (-JQ_o) + \beta_5 CC_d \times (-JQ_o) + X'_{o,d}\eta + \varepsilon_{o,d} \tag{5.2}$$

where JQ_o denotes the origin-level judicial quality, which measures the degree of contractual enforcement in origin country o. To facilitate interpretation, we negate the measure of judicial quality (-JQ) to measure *judicial weakness*. Thus, a higher value of "minus" judicial quality—(i.e.,) judicial weakness (-JQ)—indicates *weaker* institutional quality.[53]

Table 8 reports the results on the role of the judicial quality of origin countries and its interaction with credit constraints in destination US counties. First, we reconfirm the key estimation results in Table 7 such that the coefficients of the double-interaction term between the co-ethnic network and the destination-level credit constraint are all positive and statistically significant. Then, turning our attention to the role of contractual enforcement in origin countries, the coefficients of the triple-interaction term in the third row are all positive and significant, which means that the *stronger* effect of co-ethnic networks on global supply chain relationships in more credit-constrained destination counties becomes *more stronger* if origin countries have weaker judicial quality. In other words, the role of co-ethnic networks in relaxing the credit constraints of firms in a destination US county for global supply chain formation becomes even more important if the origin country has weaker institutions (i.e., lower judicial quality). By including an additional dimension of the role of judicial quality, the results reinforce the credit constraint mechanism that we hypothesized, which resonates

[52]Recall that in Section 5.2.2, we showed that common ancestry and good institutions (in origin country) are complements rather than substitutes, which is also consistent with the findings by Burchardi et al. (2019). This result is based on the interaction of log ancestry and origin country-level judicial weakness. Instead, in this section, we focus on the "triple" interaction of log ancestry, destination-level credit constraints, and origin-level judicial weakness. Therefore, we effectively investigate whether the marginal impact of immigration on global supply chain formation—which becomes stronger in destination counties with higher credit constraints—further magnifies or dampens depending on origin country's judicial weakness.

[53]Specifically, we use the measure of judicial quality from Nunn (2007), originally drawn from the "rule of law" from Kaufmann et al. (2004), which is a weighted average of a number of variables that measure individuals' assessments of the quality of the judiciary body and contract enforcement in each country.

Table 8: Heterogeneous Treatment Effect:
The Role of Judicial Quality Interacting with Credit Constraints

	100-PayDexMax		100-PayDexMin	
	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)	(3)	(4)
Log Ancestry 2010	-0.010	0.007	-0.078***	-0.042
	(0.017)	(0.022)	(0.021)	(0.027)
Log Ancestry 2010 × CC	0.287***	0.233***	0.404***	0.314***
	(0.032)	(0.044)	(0.046)	(0.048)
Log Ancestry 2010 × CC × (-JQ)	0.073***	0.091***	0.158***	0.148***
	(0.022)	(0.029)	(0.028)	(0.031)
Log Ancestry 2010 × (-JQ)	-0.049***	-0.062***	-0.099***	-0.096***
	(0.014)	(0.015)	(0.016)	(0.019)
CC × (-JQ)	0.001	-0.005*	-0.003	-0.007**
	(0.003)	(0.003)	(0.003)	(0.003)
First-stage F stat	20.9	20.9	118.0	118.0
Destination FE	✓	✓	✓	✓
Origin FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	437904	437904	437904	437904

Notes. This table presents the heterogeneous treatment effect results by including the interaction of Log Ancestry 2010 with credit constraint measures (destination level) and the measure of "minus" judicial quality (i.e., judicial weakness) (origin level). We use the same specifications as that in Table 7, where we additionally include the triple interaction among Log Ancestry 2010, credit constraints, and judicial weakness (Log Ancestry 2010 x CC x (-JQ)); the interaction between Log Ancestry 2010 and judicial weakness (Log Ancestry 2010 x (-JQ)); and the interaction between credit constraints and judicial weakness (CC x (-JQ)). The measure of credit constraints (CC) is obtained as the average credit constraint of establishments within each destination county, and the measure of judicial quality (JQ), which comes from Nunn (2007), is defined at the origin country level. To facilitate the interpretation of the coefficients, the credit constraint variables and measure of judicial quality are standardized so that the sample mean equals zero and the sample standard deviation equals one.

with the findings of Antras and Foley (2015).[54]

[54] In Appendix Table B.26, we run similar triple interaction regressions by using origin country's financial weakness instead of judicial weakness. We find similar results. That is, the *stronger* effect of co-ethnic networks on global supply chain relationships in more credit-constrained destination counties becomes *more stronger* if origin countries have weaker financial development.

6 Discussion: Aggregate and Distributional Implications

While our empirical findings move one step forward to unravel the causal, positive impacts of co-ethnic networks shaped by century-long immigration on global supply chain formation, we have not quantified the aggregate and distributional effects of immigration on the global supply chain networks in the US. For instance, if a supply chain relationship between a US county and a foreign country is arbitrary aside from conventional determinants such as transportation and production costs, then a small cost saving from co-ethnic networks can generate a huge amount of re-sorting of the global supply chain networks. In such a case, the complete removal of co-ethnic features in the US global supply chain formations could cause a mere reallocation of the global supply chain relationships with negligible overall economic efficiency. In other words, immigration may yield no net increase in global supply chain relationships in the US. Contrary to this view, however, if co-ethnic networks serve as a significant determinant of the global supply chains, then immigration might also create a net increase in global supply chains in the US.

Even though quantifying the overall economic efficiency would be a difficult task, there may be some reasons that could support the net increase of global supply chain formations as a consequence of immigration to the US. In a multi-country model of two-sided matching between a supplier and a customer, i.e., a global supply chain network model, Choi (2021) shows theoretically that a reduction in bilateral matching costs between two countries can affect unrelated third countries in a non-trivial way, but it always increases more matches in the countries experiencing the reduction in bilateral matching costs.[55] The increased number of cross-country matches from the two countries always dominates the substitution effect including both the reductions in within-country matches and the decreases in cross-country matches with the unrelated third countries.[56] In our context, because immigration to the US reduces matching and search costs between the US and a certain foreign country, it may yield a net increase in global supply chain formations, at least qualitatively, which feeds through enhanced overall economic efficiency.[57]

Next, even if the total volume of global supply chain relationships does not respond at all to immigration, which we believe is unlikely due to the above reason, it will still matter for the distribution of global supply chains across countries, which has important welfare implications and timely significance in the current debate on international fragmentation. For instance, if more Italian migrants in Napa county reduce supply chains with non-Italian firms as much as it increases supply chains with Italian ones, so that there is no positive aggregate impact of immigration on the total

[55] See Proposition 1 of Choi (2021).

[56] The sufficient statistics for welfare in each country is the total unmatched customers/suppliers and thus the welfare in the countries experiencing the reduction in bilateral matching costs unambiguously increases.

[57] Multi-country quantifiable trade models widely used in trade literature also feature a similar component (e.g., Eaton and Kortum, 2002, and many others). Reductions in any bilateral trade costs between two countries in a multi-country setup yield "gains from trade" in the involved two countries.

number of US global supply chains, our finding still has *distributional* implications. Importantly, countries can reduce economic volatility by diversifying the sources of demand and supply across countries through trade (Caselli et al., 2020), and the idea that entrepreneurs make export decisions to lower the variance of global sales has significant welfare implications (Esposito, 2022). Moreover, after the supply chain disruptions during the Covid-19 pandemic, both policymakers and academic economics literature showed renewed interest in supply chain diversification (Hyun et al., 2020).

To sum up, our empirical findings have both aggregate and distributional implications. In order to fully quantify these effects, one may need to build a quantifiable firm-to-firm trade model that features co-ethnic networks, credit frictions, and multiple countries and multiple counties in the US. Although such a full-fledged quantitative model is beyond the scope of this paper, we believe that the causal, reduced-form estimates (or elasticities) that we unraveled in this paper will be useful to better understand non-trivial efficiency and distributional consequences of immigration on the global supply chains in the US.

7 Conclusion

This article has advanced our understanding of the formation of global supply chain networks by exploring the role of immigration and its interaction with credit frictions. Our main findings are summarized as follows: (i) Co-ethnic networks formed by century-long immigration to the US causally explain the present-day structure of global supply chains, and (ii) the positive impact of immigration on global supply chain relationships becomes stronger for US counties with more credit-constrained firms, with such a stronger effect becoming even more elevated for foreign firms operating in countries with weak contract enforcement.

These findings provide new insights into and predictions for the interdependencies among global linkages of supply chains, cross-border credit provision, and immigration. First, our findings imply that the COVID-19 pandemic and the Russo-Ukrainian War may reshape the future structure of global supply chain relationships for several decades or even a century since they have already altered the pattern of migrations. Whether this prediction holds will be an important avenue for rigorous future work.

Second, the finding that co-ethnic networks mitigate credit-constraint problems in global supply chain networks indicates the tight interlinkages between trade and finance in global supply chains. In particular, this finding implies that credit provision across borders is crucial in fostering global supply chain relationships. Further research is warranted to dissect the interdependencies between global supply chains and finance, especially using global firm-to-firm trade credit provision data when such data become easily accessible to researchers interested in this topic.

Finally, by combining our work with the existing research on the consequences of supply chain networks, we can predict the far-reaching implications of immigration. For example, global supply chain networks have been shown to lead to the cross-country transmission of shocks (e.g., Boehm et al., 2019; Hyun et al., 2020; Carvalho et al., 2021). Combining it with our results, we can hypothesize that immigration from a foreign country leads to economic shock synchronization with that country. Hence, the nexus between immigration and business cycle co-movement will be another fruitful research topic that deserves special attention.

References

Acemoglu, Daron and Pablo D Azar, "Endogenous production networks," *Econometrica*, 2020, *88* (1), 33–82.

Ahn, JaeBin, "Trade finance and payment methods in international trade," in "Oxford Research Encyclopedia of Economics and Finance" 2021.

Akey, Pat and Ian Appel, "The limits of limited liability: Evidence from industrial pollution," *Journal of Finance*, 2021, *76* (1), 5–55.

Aleksynska, Mariya and Giovanni Peri, "Isolating the network effect of immigrants on trade," *World Economy*, 2014, *37* (3), 434–455.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg, "Fractionalization," *Journal of Economic Growth*, 2003, *8* (2), 155–194.

Alfaro, Laura and Maggie Xiaoyang Chen, "Surviving the global financial crisis: Foreign ownership and establishment performance," *American Economic Journal: Economic Policy*, 2012, *4* (3), 30–55.

Amiti, Mary and David E. Weinstein, "Exports and financial shocks," *Quarterly Journal of Economics*, 2011, *126* (4), 1841–1877.

Anderson, Siwan and Jean-Marie Baland, "The Economics of roscas and intrahousehold resource allocation," *Quarterly Journal of Economics*, 2002, *117* (3), 963–995.

Antras, Pol and C Fritz Foley, "Poultry in motion: a study of international trade finance practices," *Journal of Political Economy*, 2015, *123* (4), 853–901.

_ **and Elhanan Helpman**, "Global sourcing," *Journal of Political Economy*, 2004, *112* (3), 552–580.

_ **, Davin Chor, Thibault Fally, and Russell Hillberry**, "Measuring the upstreamness of production and trade flows," *American Economic Review: Papers & Proceedings*, 2012, *102* (3), 412–416.

_ **, Mihir A Desai, and C Fritz Foley**, "Multinational firms, FDI flows, and imperfect capital markets," *Quarterly Journal of Economics*, 2009, *124* (3), 1171–1219.

Avramidis, Panagiotis, George Pennacchi, Konstantinos Serfes, and Kejia Wu, "The role of regulation and bank competition in small firm financing: Evidence from the Community Reinvestment Act," *Journal of Money, Credit and Banking*, 2022.

Barnatchez, Keith, Leland Dod Crane, and Ryan Decker, "An assessment of the national establishment time series (nets) database," Technical Report 2017.

Basco, Sergi, "Financial development and the product cycle," *Journal of Economic Behavior & Organization*, 2013, *94*, 295–313.

Beck, Thorsten, "Financial development and international trade: Is there a link?," *Journal of International Economics*, 2002, *57* (1), 107–131.

Behrens, Kristian, Brahim Boualam, Julien Martin, and Florian Mayner, "Gentrification and pioneer businesses," *forthcoming at Review of Economics and Statistics*, 2022.

Bengtsson, Ola and David H Hsu, "Ethnic matching in the US venture capital market," *Journal of Business Venturing*, 2015, *30* (2), 338–354.

Bernard, Andrew B, J Bradford Jensen, Stephen J Redding, and Peter K Schott, "Firms in international trade," *Journal of Economic Perspectives*, 2007, *21* (3), 105–130.

Besley, Timothy, Stephen Coate, and Glenn Loury, "The economics of rotating savings and credit associations," *American Economic Review*, 1993, *83* (4), 792–810.

Boehm, Christoph E., Aaron Flaaen, and Nitya Pandalai-Nayar, "Input linkages and the transmission of shocks: firm-level evidence from the 2011 Tōhoku earthquake," *Review of Economics and Statistics*, 2019, *101* (1), 60–75.

Bonadio, Barthélémy, "Migrants, trade and market access," 2020.

Borisov, Alexander, Andrew Ellul, and Merih Sevilir, "Access to public capital markets and employment growth," *Journal of Financial Economics*, 2021, *141* (3), 896–918.

Burchardi, Konrad B, Thomas Chaney, and Tarek A Hassan, "Migrants, ancestors, and foreign investments," *Review of Economic Studies*, 2019, *86* (4), 1448–1486.

Cardoso, Miguel and Ananth Ramanaryanan, "Immigrants and exports: Firm-level evidence from Canada," *Canadian Journal of Economics*, 2019, *Forthcoming.*

Carluccio, Juan and Thibault Fally, "Global sourcing under imperfect capital markets," *Review of Economics and Statistics*, 2012, *94* (3), 740–763.

Carvalho, Vasco M, Makoto Nirei, Yukiko U Saito, and Alireza Tahbaz-Salehi, "Supply chain disruptions: Evidence from the great east japan earthquake," *Quarterly Journal of Economics*, 2021, *136* (2), 1255–1321.

Caselli, Francesco, Miklos Koren, Milan Lisicky, and Silvana Tenreyro, "Diversification through trade," *Quarterly Journal of Economics*, 2020, *135* (1), 449–502.

Chaney, Thomas, "Liquidity constrained exporters," *Journal of Economic Dynamics and Control*, 2016, *72*, 141–154.

Choi, Jaerim, "The global value chain under imperfect capital markets," *World Economy*, 2020, *43* (2), 484–505.

_ , "Two-sided heterogeneity, endogenous sharing, and international matching markets," *Economic Theory*, 2021, *72* (2), 473–509.

Coase, Ronald Harry, "The nature of the firm," *Economica*, 1937, *4* (16), 386–405.

Cohen, Lauren, Umit G Gurun, and Christopher Malloy, "Resident networks and corporate connections: Evidence from World War II internment camps," *Journal of Finance*, 2017, *72* (1), 207–248.

Desai, Mihir A, C Fritz Foley, and James R Hines Jr, "A multinational perspective on capital structure choice and internal capital markets," *Journal of Finance*, 2004, *59* (6), 2451–2487.

Eaton, Jonathan and Samuel Kortum, "Technology, geography, and trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Ederer, Florian and Frédéric Schneider, "Trust and promises over time," *American Economic*

Journal: Microeconomics, 2022, *14* (3), 304–320.

Esposito, Federico, "Demand risk and diversification through international trade," *Journal of International Economics*, 2022, *135*, 103562.

Fafchamps, Marcel, "Ethnicity and credit in African manufacturing," *Journal of Development Economics*, 2000, *61* (1), 205–235.

Foley, C Fritz and William R Kerr, "Ethnic innovation and US multinational firm activity," *Management Science*, 2013, *59* (7), 1529–1544.

Fukuyama, Francis, *Trust: The social virtues and the creation of prosperity*, Simon and Schuster, 1996.

Ghani, Ejaz, William R Kerr, and Christopher Stanton, "Diasporas and outsourcing: evidence from oDesk and India," *Management Science*, 2014, *60* (7), 1677–1697.

Gould, David M, "Immigrant links to the home country: empirical implications for US bilateral trade flows," *Review of Economics and Statistics*, 1994, pp. 302–316.

Gray, John V, Enno Siemsen, and Gurneeta Vasudeva, "Colocation still matters: Conformance quality and the interdependence of R&D and manufacturing in the pharmaceutical industry," *Management Science*, 2015, *61* (11), 2760–2781.

Grossman, Sanford J and Oliver D Hart, "The costs and benefits of ownership: A theory of vertical and lateral integration," *Journal of Political Economy*, 1986, *94* (4), 691–719.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "The role of social capital in financial development," *American Economic Review*, 2004, *94* (3), 526–556.

Hart, Oliver and John Moore, "Incomplete contracts and renegotiation," *Econometrica: Journal of the Econometric Society*, 1988, pp. 755–785.

— **and** — , "Property rights and the nature of the firm," *Journal of Political Economy*, 1990, *98* (6), 1119–1158.

Head, Keith and John Ries, "Immigration and trade creation: econometric evidence from Canada," *Canadian Journal of Economics*, 1998, pp. 47–62.

Hyun, Jay and Ryan Kim, "Spillovers through multimarket firms: The uniform product replacement channel," 2022. working paper.

— **, Daisoon Kim, and Seung-Ryong Shin**, "The role of global connectedness and market power in crises: Firm-level evidence from the COVID-19 pandemic," *Covid Economics*, 2020, *49* (1), 148–171.

— **, Ziho Park, and Vladimir Smirnyagin**, "Shock propagation within multisector firms," 2022. working paper.

Javorcik, Beata S, Çağlar Özden, Mariana Spatareanu, and Cristina Neagu, "Migrant networks and foreign direct investment," *Journal of Development Economics*, 2011, *94* (2), 231–241.

Jensen, Eric B., Renuka Bhaskar, and Melissa Scopilliti, "Demographic analysis 2010: Estimates of coverage of the foreign-born population in the American Community Survey," Working Paper POP-WP103, U.S. Census 2015.

Ju, Jiandong and Shang-Jin Wei, "Domestic institutions and the bypass effect of financial globalization," *American Economic Journal: Economic Policy*, 2010, *2* (4), 173–204.

_ **and** _ , "When is quality of financial system a source of comparative advantage?," *Journal of International Economics*, 2011, *84* (2), 178–187.

Kallberg, Jarl G and Gregory F Udell, "The value of private sector business credit information sharing: the US case," *Journal of Banking & Finance*, 2003, *27* (3), 449–469.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl, "Trust and social collateral," *Quarterly Journal of Economics*, 2009, *124* (3), 1307–1361.

Karlan, Dean S, "Using experimental economics to measure social capital and predict financial decisions," *American Economic Review*, 2005, *95* (5), 1688–1699.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi, "Governance matters III: Governance indicators for 1996, 1998, 2000, and 2002," *World Bank Economic Review*, 2004, *18* (2), 253–287.

Kim, Se-Jik and Hyun Song Shin, "Sustaining production chains through financial linkages," *American Economic Review*, 2012, *102* (3), 402–406.

Kletzer, Kenneth and Pranab Bardhan, "Credit markets and patterns of international trade," *Journal of Development Economics*, 1987, *27* (1-2), 57–70.

Kong, Dongmin, Yue Pan, Gary Gang Tian, and Pengdong Zhang, "CEOs' hometown connections and access to trade credit: Evidence from China," *Journal of Corporate Finance*, 2020, *62*, 101574.

Kopytov, Alexandr, Bineet Mishra, Kristoffer Nimark, and Mathieu Taschereau-Dumouchel, "Endogenous production networks under supply chain uncertainty," 2022. working paper.

Levine, Ross, Chen Lin, and Wensi Xie, "Corporate resilience to banking crises: The roles of trust and trade credit," *Journal of Financial and Quantitative Analysis*, 2018, *53* (4), 1441–1477.

Light, Ivan and Michelle Pham, "Beyond creditworthy: Microcredit and informal credit in the United States," *Journal of Developmental Entrepreneurship*, 1998, *3* (1), 35.

_ **, Im Jung Kwuon, and Deng Zhong**, "Korean rotating credit associations in Los Angeles," *Amerasia Journal*, 1990, *16* (2), 35–54.

Lim, Kevin, "Endogenous production networks and the business cycle," 2018. working paper.

Manova, Kalina, "Credit constraints, heterogeneous firms, and international trade," *Review of Economic Studies*, 2013, *80* (2), 711–744.

_ **, Shang-Jin Wei, and Zhiwei Zhang**, "Firm exports and multinational activity under credit constraints," *Review of Economics and Statistics*, 2015, *97* (3), 574–588.

Matsuyama, Kiminori, "Credit market imperfections and patterns of international trade and capital flows," *Journal of the European Economic Association*, 2005, *3* (2-3), 714–723.

Min, Pyong Gap, "Korean immigrants in Los Angeles 1," in Ivan Light and Parminder Bhachu, eds., *Immigration and Entrepreneurship*, Routledge, 1993, pp. 185 – 204.

Niepmann, Friederike and Tim Schmidt-Eisenlohr, "No guarantees, no trade: How banks affect export patterns," *Journal of International Economics*, 2017, *108*, 338–350.

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

Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter, "Plants in space," *Mimeo*, 2022.

Parsons, Christopher and Pierre-Louis Vézina, "Migrant networks and trade: The Vietnamese boat people as a natural experiment," *Economic Journal*, 2018, *128* (612), F210–F234.

Putnam, Robert D et al., *Bowling alone: The collapse and revival of American community*, Simon and schuster, 2000.

Rajan, Raghuram G. and Luigi Zingales, "Financial dependence and growth," *American Economic Review*, 1998, *88* (3), 559–586.

Rauch, James E, "Business and social networks in international trade," *Journal of Economic Literature*, 2001, *39* (4), 1177–1203.

_ **and Vitor Trindade**, "Ethnic Chinese networks in international trade," *Review of Economics and Statistics*, 2002, *84* (1), 116–130.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, "Diverging trends in national and local concentration," *NBER Macroeconomics Annual*, 2021, *35* (1), 115–150.

Taschereau-Dumouchel, Mathieu, "Cascades and fluctuations in an economy with an endogenous production network," 2022. working paper.

Williamson, Oliver E, *Markets and hierarchies: analysis and antitrust implications: a study in the economics of internal organization*, NY: Free Press, 1975.

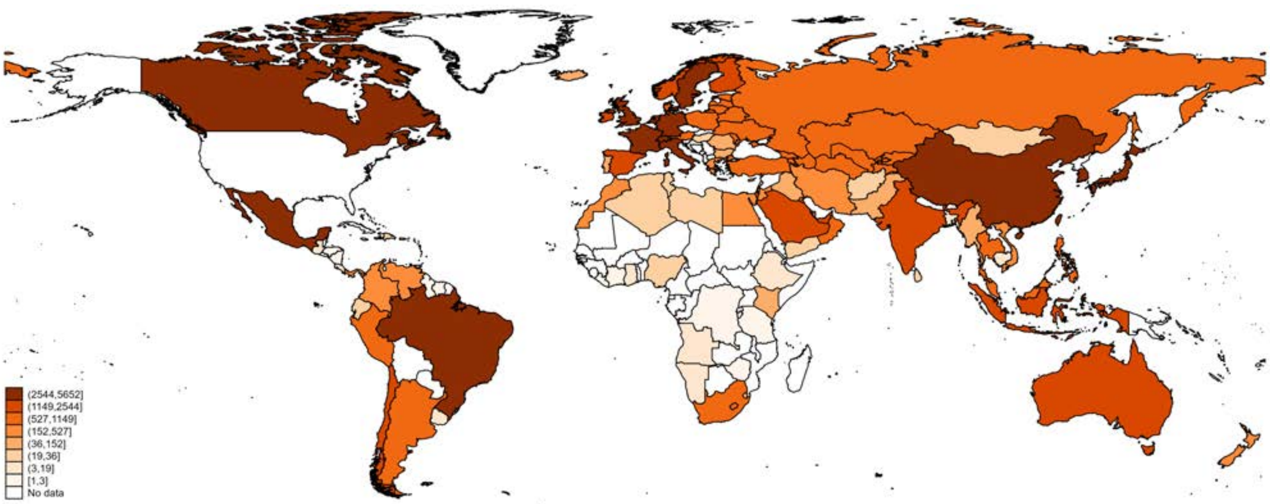
_ , *The Economic Intstitutions of Capitalism*, Simon and Schuster, 1985.

Wu, Wenfeng, Michael Firth, and Oliver M Rui, "Trust and the provision of trade credit," *Journal of Banking & Finance*, 2014, *39*, 146–159.

Xie, Fei, Bohui Zhang, and Wenrui Zhang, "Trust, incomplete contracting, and corporate innovation," *Management Science*, 2022, *68* (5), 3419–3443.

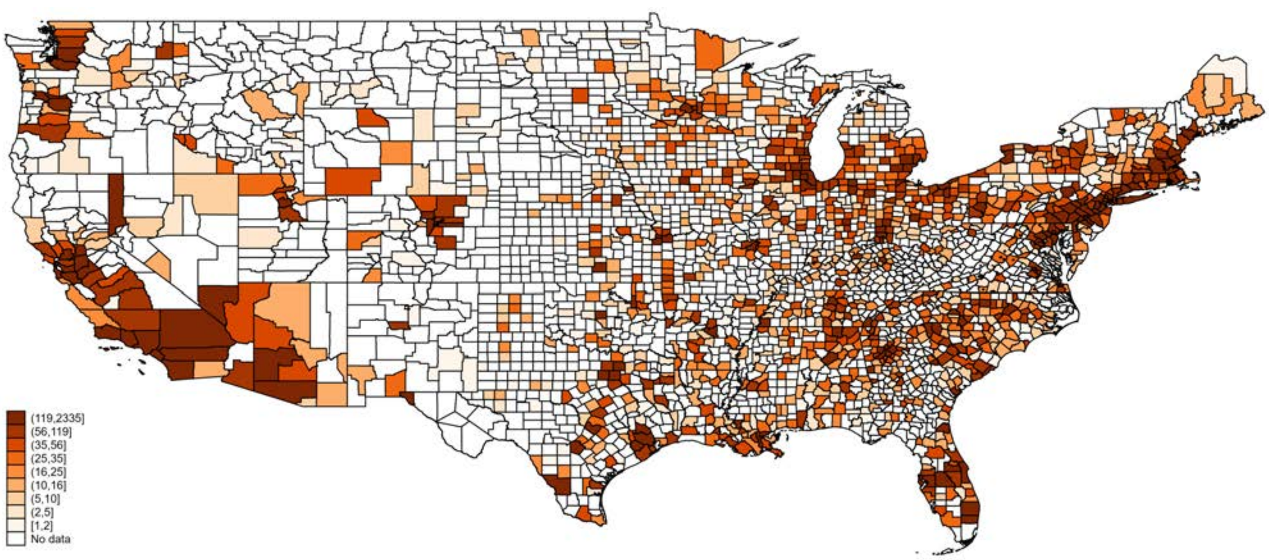
Appendix A Additional Figures

Figure A.1: Global Supply Chain Linkages by Origin Country



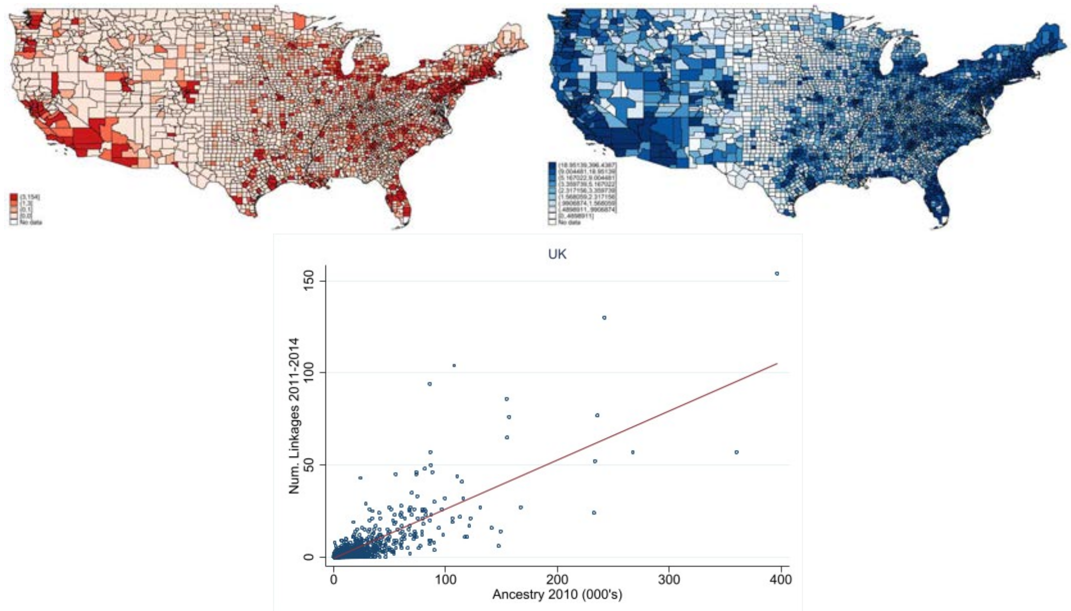
Notes. The world map shows the origin-country variation in the global supply chain linkages with US firms. A darker red indicates that a country has more connections with US suppliers and/or US customers; the figure shows a significant variation in the total number of global supply chain relationships with the US across foreign countries.

Figure A.2: Global Supply Chain Linkages by Destination County



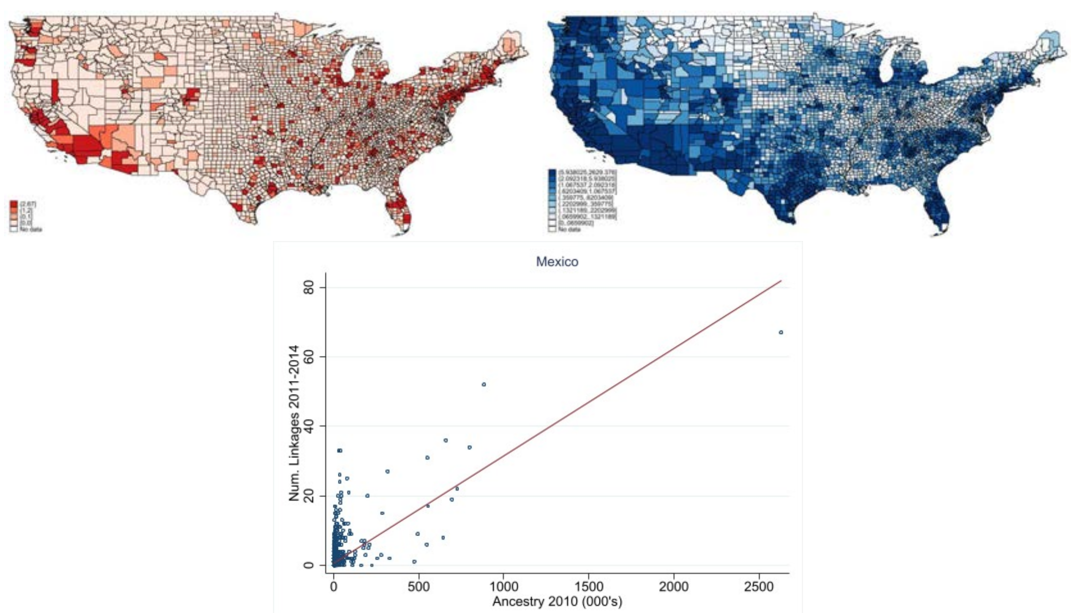
Notes. The US map shows the county variation in the global supply chain linkages within the US. A darker red indicates that a county has more connections with foreign suppliers and/or foreign customers. The figure shows a significant variation in the total number of global supply chain relationships across counties within the US.

Figure A.3: Global Supply Chain Linkages and Ancestry: United Kingdom



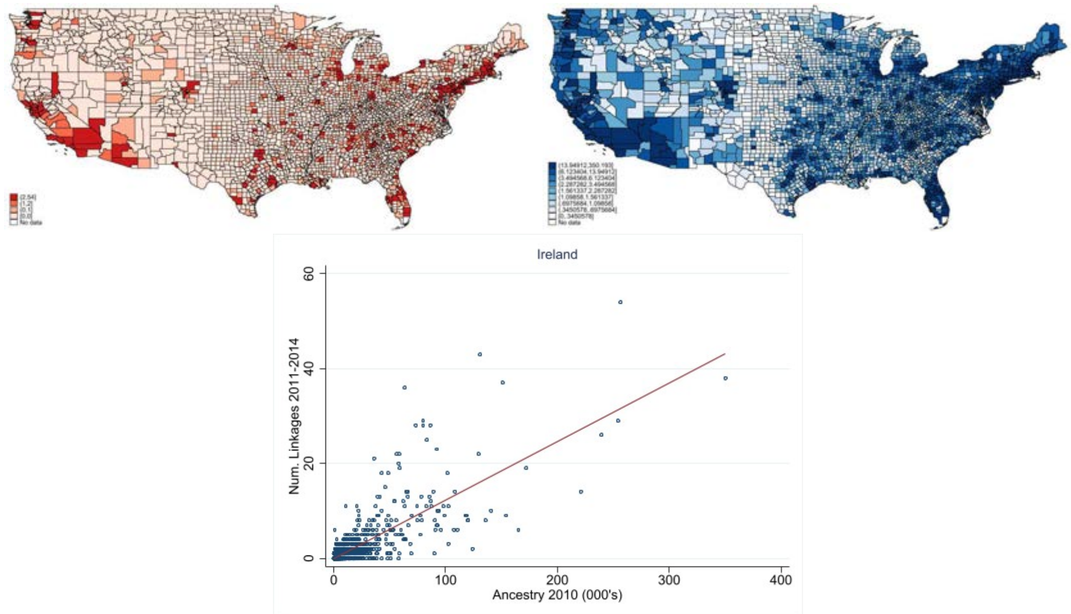
Notes. The figure is restricted to one origin country: United Kingdom. The first red map shows the county variation in the global supply chain linkages with British firms within the US. The second blue map shows the county variation in British ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of British ancestry and the number of global supply chain linkages with British firms across US counties.

Figure A.4: Global Supply Chain Linkages and Ancestry: Mexico



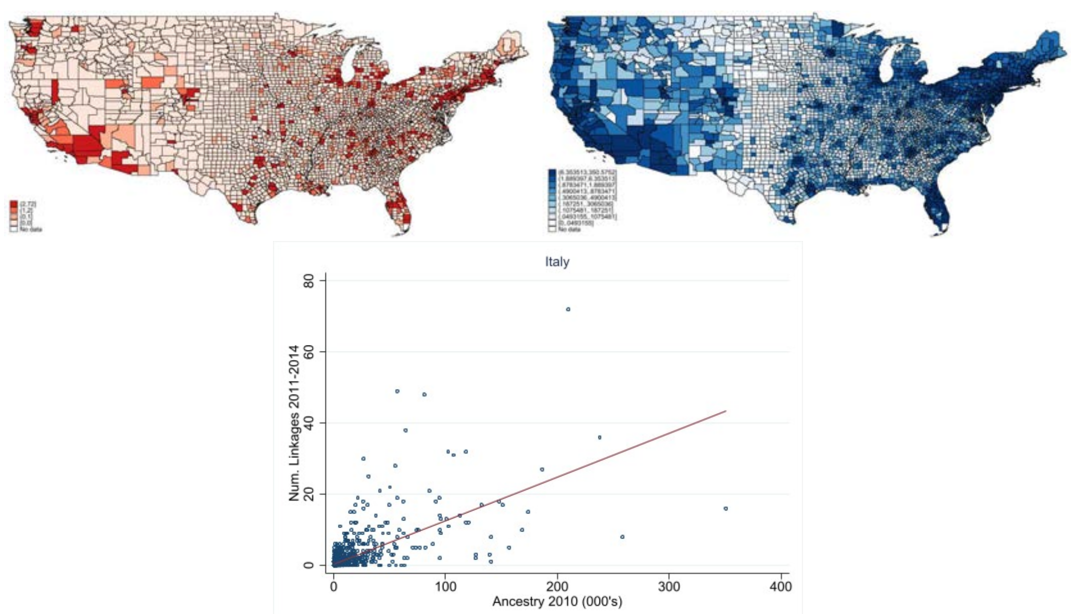
Notes. The figure is restricted to one origin country: Mexico. The first red map shows the county variation in the global supply chain linkages with Mexican firms within the US. The second blue map shows the county variation in Mexican ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of Mexican ancestry and the number of global supply chain linkages with Mexican firms across US counties.

Figure A.5: Global Supply Chain Linkages and Ancestry: Ireland



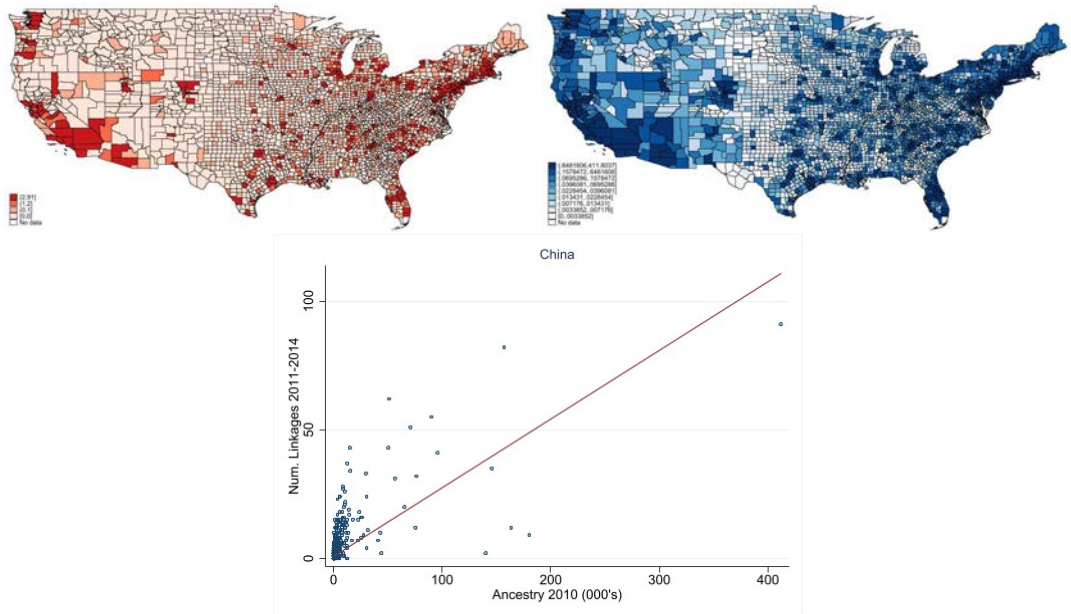
Notes. The figure is restricted to one origin country: Ireland. The first red map shows the county variation in the global supply chain linkages with Irish firms within the US. The second blue map shows the county variation in Irish ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of Irish ancestry and the number of global supply chain linkages with Irish firms across US counties.

Figure A.6: Global Supply Chain Linkages and Ancestry: Italy



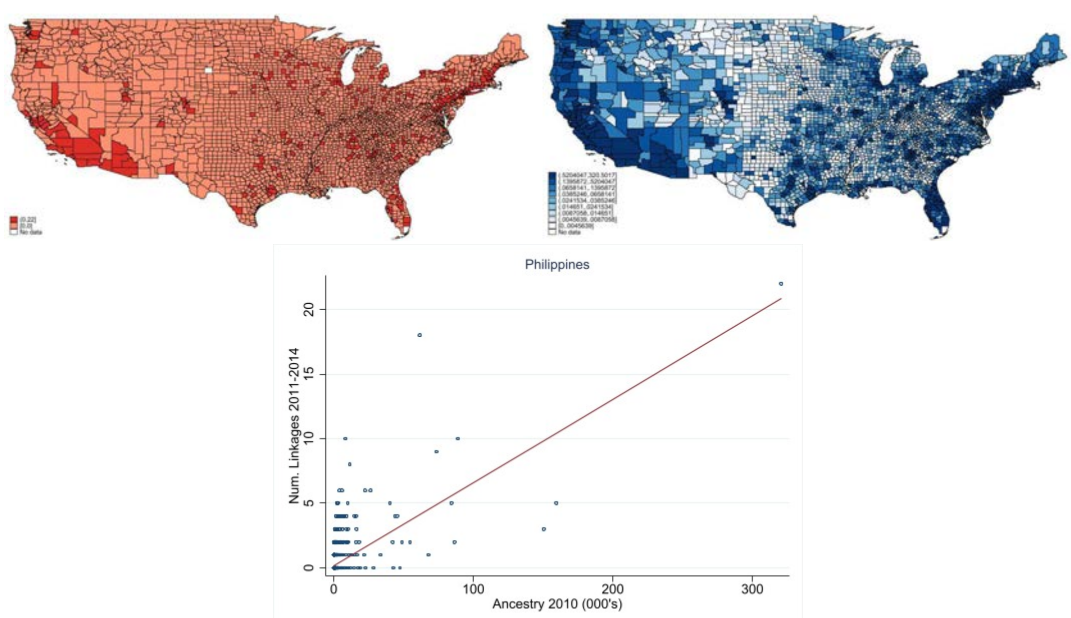
Notes. The figure is restricted to one origin country: Italy. The first red map shows the county variation in the global supply chain linkages with Italian firms within the US. The second blue map shows the county variation in Italian ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of Italian ancestry and the number of global supply chain linkages with Italian firms across US counties.

Figure A.7: Global Supply Chain Linkages and Ancestry: China



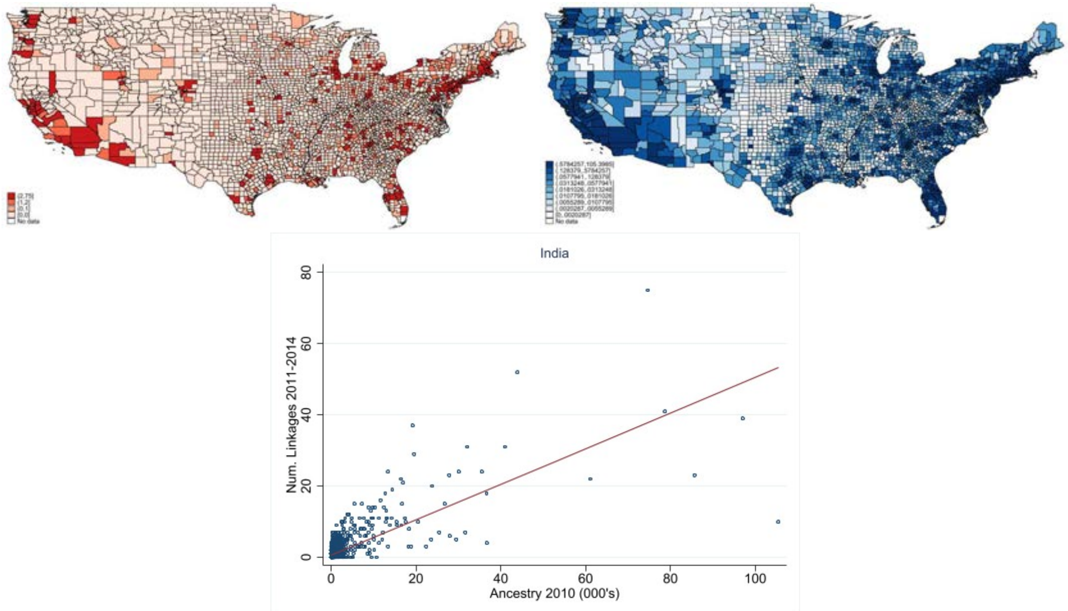
Notes. The figure is restricted to one origin country: China. The first red map shows the county variation in the global supply chain linkages with Chinese firms within the US. The second blue map shows the county variation in Chinese ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of Chinese ancestry and the number of global supply chain linkages with Chinese firms across US counties.

Figure A.8: Global Supply Chain Linkages and Ancestry: Philippines



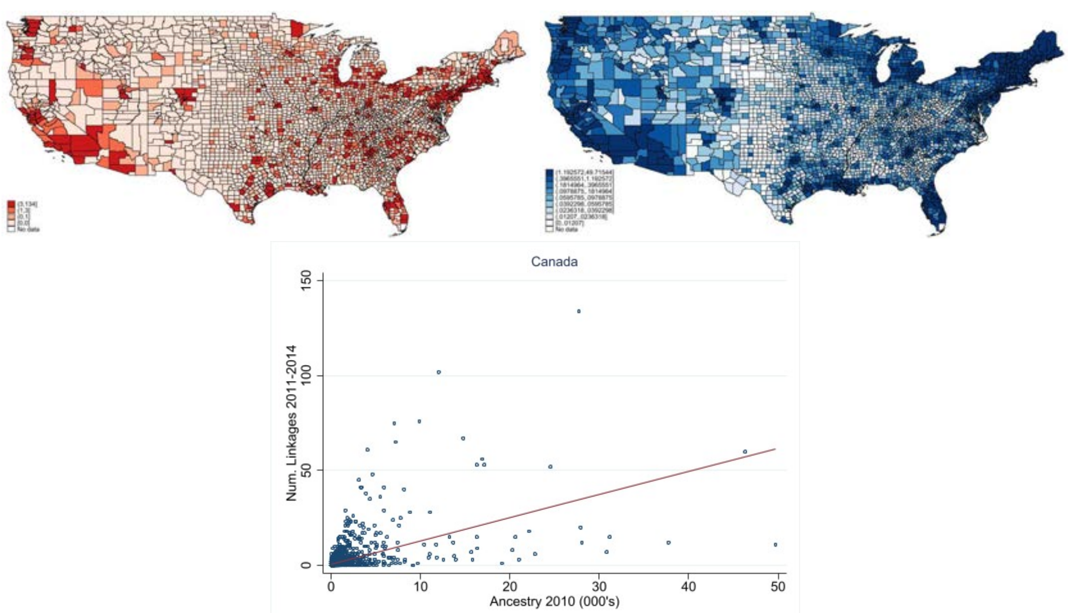
Notes. The figure is restricted to one origin country: Philippines. The first red map shows the county variation in the global supply chain linkages with Filipino firms within the US. The second blue map shows the county variation in Filipino ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of Filipino ancestry and the number of global supply chain linkages with Filipino firms across US counties.

Figure A.9: Global Supply Chain Linkages and Ancestry: India



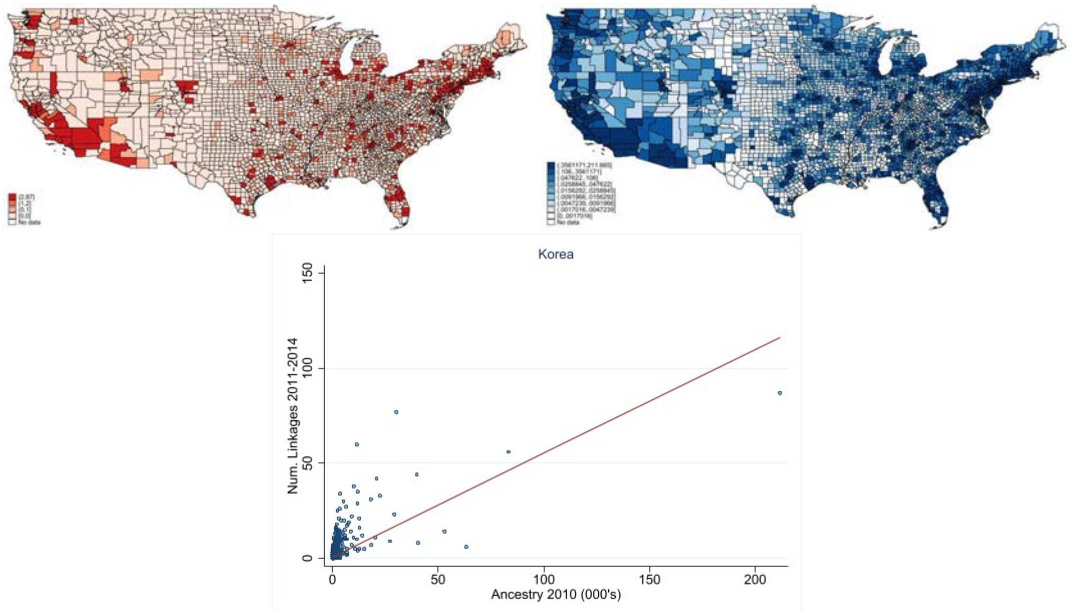
Notes. The figure is restricted to one origin country: India. The first red map shows the county variation in the global supply chain linkages with Indian firms within the US. The second blue map shows the county variation in Indian ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of Indian ancestry and the number of global supply chain linkages with Indian firms across US counties.

Figure A.10: Global Supply Chain Linkages and Ancestry: Canada



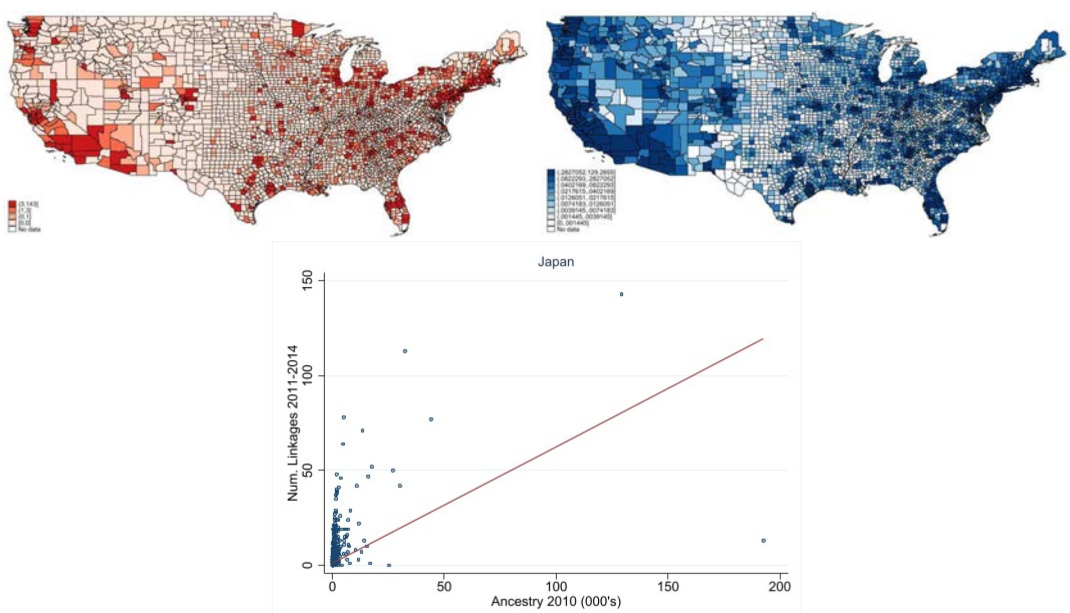
Notes. The figure is restricted to one origin country: Canada. The first red map shows the county variation in the global supply chain linkages with Canadian firms within the US. The second blue map shows the county variation in Canadian ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of Canadian ancestry and the number of global supply chain linkages with Canadian firms across US counties.

Figure A.11: Global Supply Chain Linkages and Ancestry: South Korea



Notes. The figure is restricted to one origin country: South Korea. The first red map shows the county variation in the global supply chain linkages with Korean firms within the US. The second blue map shows the county variation in Korean ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of Korean ancestry and the number of global supply chain linkages with Korean firms across US counties.

Figure A.12: Global Supply Chain Linkages and Ancestry: Japan



Notes. The figure is restricted to one origin country: Japan. The first red map shows the county variation in the global supply chain linkages with Japanese firms within the US. The second blue map shows the county variation in Japanese ancestry within the US. The third figure shows a scatter plot (and its linear fit) between the number of Japanese ancestry and the number of global supply chain linkages with Japanese firms across US counties.

Appendix B Additional Tables

Table B.1: Top 30 Largest Origin Countries in the US

Panel A. Rank Based on # of Ancestry				**Panel B. Rank Based on # of GSC Links**			
Rank	Country	# of Ancestry (000's)	Share (%)	Rank	Country	# of GSC Links	Share (%)
1	Germany	35,314	18.2	1	United Kingdom	1,086	8.7
2	United Kingdom	28,150	14.5	2	Japan	847	6.8
3	Mexico	22,904	11.8	3	Germany	833	6.7
4	Ireland	22,240	11.5	4	Canada	833	6.7
5	Italy	13,759	7.1	5	France	742	5.9
6	Poland	6,833	3.5	6	Netherlands	553	4.4
7	France	5,810	3.0	7	China	540	4.3
8	USSR	5,003	2.6	8	Switzerland	510	4.1
9	China	3,580	1.8	9	Republic of Korea	486	3.9
10	Norway	3,107	1.6	10	Italy	387	3.1
11	Sweden	3,002	1.6	11	Sweden	387	3.1
12	Spain	2,888	1.5	12	Australia	378	3.0
13	Netherlands	2,880	1.5	13	Israel	352	2.8
14	Philippines	2,729	1.4	14	India	340	2.7
15	India	2,433	1.3	15	Ireland	300	2.4
16	Canada	2,130	1.1	16	Mexico	276	2.2
17	Cuba	1,552	0.8	17	Brazil	264	2.1
18	Republic of Korea	1,529	0.8	18	Singapore	250	2.0
19	Czechoslovakia	1,507	0.8	19	Spain	204	1.6
20	El Salvador	1,499	0.8	20	Saudi Arabia	178	1.4
21	Vietnam	1,441	0.7	21	Finland	178	1.4
22	Dominican Republic	1,418	0.7	22	Norway	155	1.2
23	Japan	1,144	0.6	23	Indonesia	151	1.2
24	Colombia	1,112	0.6	24	Chile	140	1.1
25	Greece	1,105	0.6	25	USSR	124	1.0
26	Portugal	1,082	0.6	26	United Arab Emirates	117	0.9
27	Hungary	990	0.5	27	South Africa	116	0.9
28	Guatemala	964	0.5	28	Denmark	109	0.9
29	Yugoslavia	899	0.5	29	Turkey	107	0.9
30	Denmark	887	0.5	30	Malaysia	99	0.8

Notes. The table reports the top 30 largest origin countries in the US. The ranking in Panel A (Panel B, respectively) is based on 30 foreign countries with the largest ancestry in 2010 (GSC linkages in 2011-2014, respectively). The top 30 origin countries account for 92.9% (88.2%, respectively) of the total ancestry groups (total GSC linkages, respectively) in the US. 17 countries appear in both Panels: Germany, United Kingdom, Mexico, Ireland, Italy, France, USSR, China, Norway, Sweden, Spain, Netherlands, India, Canada, Republic of Korea, Japan, and Denmark.

Table B.2: Relationship between Ancestry and GSC Linkages: Origin Country Level

	Log N.Link
	(1)
Log Ancestry 2010	0.269**
	(0.110)
Log Country Population 2010	0.722***
	(0.124)
Log GDP per Capita 2010	1.251***
	(0.177)
Continent FE	✓
Controls	✓
Observations	99

Notes. This table regresses the log of the number of global supply chain linkages with the US (2011-2014) on log number of ancestry (2010) at the origin country level. We control for the log of country population and log of GDP per capita, as well as distance and latitude difference with US, measure of ethnic fractionalization (Alesina et al., 2003), FDI indicator, and continent fixed effects. Regression is weighted by country population. Robust standard errors are in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3: The Impact of Ancestry Composition on Global Supply Chain Linkages: OLS Regressions

Panel A	I(N.Supp>0)				
	(1)	(2)	(3)	(4)	(5)
Log Ancestry 2010	0.196***	0.184***	0.183***	0.182***	0.195***
	(0.014)	(0.016)	(0.016)	(0.017)	(0.017)
Log Distance	0.012	0.045	0.022	0.051	-0.008
	(0.015)	(0.034)	(0.040)	(0.051)	(0.051)
Destination FE	✓	-	-	-	-
Origin FE	✓	-	-	-	-
Destination x Continent FE	-	✓	✓	✓	✓
Origin x Census Division FE	-	✓	✓	✓	-
3rd order poly in dist and lat	-	-	✓	✓	-
Agricultural Similarity	-	-	-	✓	-
Origin x State FE	-	-	-	-	✓
Observations	612495	612495	612495	459150	612300

Panel B	I(N.Cust>0)				
	(6)	(7)	(8)	(9)	(10)
Log Ancestry 2010	0.189***	0.173***	0.173***	0.172***	0.184***
	(0.014)	(0.016)	(0.016)	(0.017)	(0.017)
Log Distance	0.014	0.033	0.020	0.038	-0.027
	(0.018)	(0.038)	(0.042)	(0.056)	(0.071)
Destination FE	✓	-	-	-	-
Origin FE	✓	-	-	-	-
Destination x Continent FE	-	✓	✓	✓	✓
Origin x Census Division FE	-	✓	✓	✓	-
3rd order poly in dist and lat	-	-	✓	✓	-
Agricultural Similarity	-	-	-	✓	-
Origin x State FE	-	-	-	-	✓
Observations	612495	612495	612495	459150	612300

Notes. This table presents the OLS regression results that correspond to the IV regression results in Table 2. Columns (1)-(5) correspond to columns (1), (3), (5), (6), and (7) in Table 2; columns (6)-(10) correspond to columns (8), (10), (12), (13), and (14) in Table 2.

Table B.4: Using Linkage Dummy

	I(N.Link>0)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Ancestry 2010	0.228***	0.194***	0.198***	0.193***	0.198***	0.198***	0.203***
	(0.019)	(0.026)	(0.019)	(0.020)	(0.019)	(0.022)	(0.020)
Log Distance	0.017	0.014	0.039	0.038	0.018	0.048	-0.033
	(0.019)	(0.018)	(0.039)	(0.039)	(0.043)	(0.055)	(0.063)
First-stage F stat	11.0	2448.0	162.2	195.4	158.1	102.8	186.2
Destination FE	✓	✓	-	-	-	-	-
Origin FE	✓	✓	-	-	-	-	-
Principal Components	-	✓	✓	✓	✓	✓	✓
Destination x Continent FE	-	-	✓	✓	✓	✓	✓
Origin x Census Division FE	-	-	✓	✓	✓	✓	-
$I^{2010}_{o,-r(d)}(I^{2010}_{-c(o),d}/I^{2010}_{-c(o)})$	-	-	-	✓	-	-	-
3rd order poly in dist and lat	-	-	-	-	✓	✓	-
Agricultural Similarity	-	-	-	-	-	✓	-
Origin x State FE	-	-	-	-	-	-	✓
Observations	612495	612495	612495	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we use a union of supplier-firm and customer-firm linkages, I(N.Link>0), as the dependent variable. I(N.Link>0) is formally defined as a dummy variable that equals one if any firm whose trade-engaging establishments located in destination d has *either* supplier firms *or* customer firms headquartered in origin o. Columns (1)-(7) use the specifications in columns (1)-(7) in Table 2, where only the dependent variable is replaced with the linkage dummy.

Table B.5: Intensive Margin: Log Number of Linkages

Panel A	Log N.Supp						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Ancestry 2010	0.309***	0.213***	0.253***	0.318***	0.254***	0.248***	0.262***
	(0.110)	(0.067)	(0.074)	(0.062)	(0.074)	(0.091)	(0.075)
First-stage F stat	10.2	84.3	26.8	26.1	26.3	21.4	22.7
Destination FE	✓	✓	-	-	-	-	-
Origin FE	✓	✓	-	-	-	-	-
Principal Components	-	✓	✓	✓	✓	✓	✓
Destination x Continent FE	-	-	✓	✓	✓	✓	✓
Origin x Census Division FE	-	-	✓	✓	✓	✓	-
$I^{2010}_{o,-r(d)}(I^{2010}_{-c(o),d}/I^{2010}_{-c(o)})$	-	-	-	✓	-	-	-
3rd order poly in dist and lat	-	-	-	-	✓	✓	-
Agricultural Similarity	-	-	-	-	-	✓	-
Origin x State FE	-	-	-	-	-	-	✓
Observations	20385	20385	20385	20385	20385	18949	20340

Panel B	Log N.Cust						
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Log Ancestry 2010	0.341***	0.204***	0.255***	0.279***	0.254***	0.279***	0.265***
	(0.094)	(0.054)	(0.069)	(0.055)	(0.068)	(0.085)	(0.072)
First-stage F stat	14.9	67.7	16.2	31.3	17.0	21.5	18.4
Destination FE	✓	✓	-	-	-	-	-
Origin FE	✓	✓	-	-	-	-	-
Principal Components	-	✓	✓	✓	✓	✓	✓
Destination x Continent FE	-	-	✓	✓	✓	✓	✓
Origin x Census Division FE	-	-	✓	✓	✓	✓	-
$I^{2010}_{o,-r(d)}(I^{2010}_{-c(o),d}/I^{2010}_{-c(o)})$	-	-	-	✓	-	-	-
3rd order poly in dist and lat	-	-	-	-	✓	✓	-
Agricultural Similarity	-	-	-	-	-	✓	-
Origin x State FE	-	-	-	-	-	-	✓
Observations	22968	22968	22968	22968	22968	20551	22931

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we consider the intensive margin of global supply chain linkages as dependent variables. Specifically, we use the log of the number of supplier linkages (Log N.Supp) and customer linkages (Log N.Cust), respectively, as dependent variables. Columns (1)-(7) use the specifications in columns (1)-(7) in Table 2, where only the dependent variable is replaced with Log N.Supp; columns (8)-(14) use the specifications in columns (8)-(14) in Table 2, where only the dependent variable is replaced with Log N.Cust.

Table B.6: Strictly Restricting US Firms to Those Who Have Headquarters Located in the US

	I(N.Supp>0)				I(N.Cust>0)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Ancestry 2010	0.178***	0.177***	0.191***	0.188***	0.161***	0.161***	0.167***	0.167***
	(0.019)	(0.019)	(0.022)	(0.020)	(0.019)	(0.019)	(0.023)	(0.020)
First-stage F stat	76.3	75.9	60.2	84.8	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	612495	612495	459150	612300	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we strictly restrict US firms to those who have headquarters located in the US. Therefore, in this exercise, global supply chain linkages capture relationship between foreign firms whose headquarters are located outside the US and US firms whose headquarters are located in the US. This exercise strictly rules out any potential supplier-customer linkages that might capture within-multinational relationships between headquarters and their subsidiaries. Columns (1)-(4) use the specification in columns (1)-(4) in Table 3; columns (5)-(8) use the specifications in columns (5)-(8) in Table 3.

Table B.7: Restricting Establishments:
(i) Single-establishment Firms, (ii) Headquarter Establishments

Panel A	Single-establishment Firms							
	I(N.Supp>0)				I(N.Cust>0)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Ancestry 2010	0.094***	0.094***	0.103***	0.103***	0.108***	0.108***	0.113***	0.115***
	(0.022)	(0.022)	(0.024)	(0.022)	(0.020)	(0.020)	(0.022)	(0.021)
First-stage F stat	76.3	75.9	60.2	84.8	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	612495	612495	459150	612300	612495	612495	459150	612300

Panel B	Headquarter Establishments							
	I(N.Supp>0)				I(N.Cust>0)			
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Log Ancestry 2010	0.142***	0.142***	0.150***	0.154***	0.151***	0.151***	0.161***	0.163***
	(0.027)	(0.027)	(0.029)	(0.027)	(0.023)	(0.023)	(0.027)	(0.024)
First-stage F stat	76.3	75.9	60.2	84.8	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	612495	612495	459150	612300	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where Panel A measures firms' US county location only using (trade-engaging) single-establishment firms and Panel B measures location using (trade-engaging) headquarter establishments. Columns (1)-(4) in Panel A and columns (9)-(12) in Panel B use the specification in columns (1)-(4) in Table 3; columns (5)-(8) in Panel A and columns (13)-(16) in Panel B use the specifications in columns (5)-(8) in Table 3.

Table B.8: Using Ancestry Compositions Prior to 2010

	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)	(3)	(4)	(5)	(6)
Log Ancestry 2000	0.223***	0.205***				
	(0.023)	(0.023)				
Log Ancestry 1990			0.238***	0.219***		
			(0.026)	(0.026)		
Log Ancestry 1980					0.261***	0.240***
					(0.032)	(0.032)
First-stage F stat	10.9	10.9	10.7	10.7	12.1	12.1
Destination x Continent FE	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	✓	✓	✓
Observations	612495	612495	612495	612495	612495	612495

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we use the log of Ancestry in 2000, 1990, and 1980, respectively. Columns (1), (3), (5) and columns (2), (4), (6) use the specifications in columns (3) and (10) in Table 2, respectively, except that the principal components are excluded. We additionally exclude IVs that are constructed by using waves 1990 and 2000.

Table B.9: Restricting Global Supply Chains to Newly Formed Linkages After 2010

	I(N.Supp>0)				I(N.Cust>0)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Ancestry 2010	0.164***	0.163***	0.173***	0.175***	0.167***	0.167***	0.175***	0.176***
	(0.027)	(0.027)	(0.026)	(0.028)	(0.019)	(0.019)	(0.021)	(0.020)
First-stage F stat	76.3	75.9	60.2	84.8	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	612495	612495	459150	612300	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we measure linkages based on global supply chain connections that did not exist in 2010 but were newly formed between 2011 and 2014. Columns (1)-(4) use the specification in columns (1)-(4) in Table 3; columns (5)-(8) use the specifications in columns (5)-(8) in Table 3.

Table B.10: Excluding US Counties in West Coast and Countries in Asia

	I(N.Supp>0)				I(N.Cust>0)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Ancestry 2010	0.212***	0.212***	0.222***	0.226***	0.205***	0.205***	0.210***	0.217***
	(0.026)	(0.026)	(0.028)	(0.027)	(0.023)	(0.023)	(0.026)	(0.023)
First-stage F stat	81.4	79.5	64.5	90.1	81.4	79.5	64.5	90.1
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	418035	418035	311638	417880	418035	418035	311638	417880

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we exclude US counties in the West census region (i.e., Census Region 4) and countries in Asia. Columns (1)-(4) use the specification in columns (1)-(4) in Table 3; columns (5)-(8) use the specifications in columns (5)-(8) in Table 3.

Table B.11: Dropping Top Five Ancestry Origin Countries

Panel A	I(N.Supp>0)					
	(1)	(2)	(3)	(4)	(5)	(6)
Log Ancestry 2010	0.176***	0.177***	0.149***	0.177***	0.176***	0.139***
	(0.021)	(0.019)	(0.027)	(0.019)	(0.019)	(0.029)
First-stage F stat	78.3	80.1	33.6	71.1	73.6	63.9
Principal Components	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	✓	✓	✓
Drop	DEU	GBR	MEX	IRL	ITA	All
Observations	609354	609354	609354	609354	609354	596790
Panel B	I(N.Cust>0)					
	(7)	(8)	(9)	(10)	(11)	(12)
Log Ancestry 2010	0.161***	0.160***	0.136***	0.159***	0.155***	0.127***
	(0.020)	(0.019)	(0.027)	(0.019)	(0.019)	(0.029)
First-stage F stat	78.3	80.1	33.6	71.1	73.6	63.9
Principal Components	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	✓	✓	✓
Drop	DEU	GBR	MEX	IRL	ITA	All
Observations	609354	609354	609354	609354	609354	596790

Notes. This table presents coefficient estimates from IV regressions of Equation (3.1), where we exclude top five ancestry origin countries (based on 2010) in the regressions: Germany, United Kingdom, Mexico, Ireland, and Italy. Specifically, columns (1)-(5) and (7)-(11) drop countries one-by-one and columns (6) and (12) jointly drops all countries. Columns (1)-(6) use specification in column (1) in Table 3; columns (7)-(12) use specifications in column (5) in Table 3.

Table B.12: Dropping Canada and Top Five Asian Countries in terms of Ancestry Origin

Panel A	I(N.Supp>0)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Ancestry 2010	0.185***	0.182***	0.180***	0.170***	0.180***	0.178***	0.180***
	(0.020)	(0.020)	(0.020)	(0.017)	(0.019)	(0.019)	(0.018)
First-stage F stat	78.3	77.2	77.4	75.0	76.2	73.9	78.7
Principal Components	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	✓	✓	✓	✓
Drop	CHN	PHL	IND	CAN	KOR	JPN	All
Observations	609354	609354	609354	609354	609354	609354	593649
Panel B	I(N.Cust>0)						
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Log Ancestry 2010	0.168***	0.165***	0.163***	0.154***	0.164***	0.163***	0.168***
	(0.018)	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)	(0.017)
First-stage F stat	78.3	77.2	77.4	75.0	76.2	73.9	78.7
Principal Components	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	✓	✓	✓	✓
Drop	CHN	PHL	IND	CAN	KOR	JPN	All
Observations	609354	609354	609354	609354	609354	609354	593649

Notes. This table presents coefficient estimates from IV regressions of Equation (3.1), where we exclude Canada and top five Asian countries in terms of ancestry origin (based on 2010) in the regressions: China, Philippines, India, Canada, South Korea, and Japan (ordered based on rank). Specifically, columns (1)-(6) and (8)-(13) drop countries one-by-one and columns (7) and (14) jointly drops all countries. Columns (1)-(7) use specification in column (1) in Table 3; columns (8)-(14) use specifications in column (5) in Table 3.

Table B.13: Dropping Top Five GSC Partner Countries

Panel A	I(N.Supp>0)					
	(1)	(2)	(3)	(4)	(5)	(6)
Log Ancestry 2010	0.177***	0.178***	0.170***	0.176***	0.180***	0.157***
	(0.019)	(0.019)	(0.017)	(0.021)	(0.019)	(0.019)
First-stage F stat	80.1	73.9	75.0	78.3	78.6	109.2
Principal Components	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	✓	✓	✓
Drop	GBR	JPN	CAN	DEU	FRA	All
Observations	609354	609354	609354	609354	609354	596790
Panel B	I(N.Cust>0)					
	(7)	(8)	(9)	(10)	(11)	(12)
Log Ancestry 2010	0.160***	0.163***	0.154***	0.161***	0.163***	0.142***
	(0.019)	(0.019)	(0.018)	(0.020)	(0.019)	(0.019)
First-stage F stat	80.1	73.9	75.0	78.3	78.6	109.2
Principal Components	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	✓	✓	✓
Drop	GBR	JPN	CAN	DEU	FRA	All
Observations	609354	609354	609354	609354	609354	596790

Notes. This table presents coefficient estimates from IV regressions of Equation (3.1), where we exclude top five global supply chain partner countries (based on 2011-2014) in the regressions: United Kingdom, Japan, Canada, Germany, and France. Specifically, columns (1)-(5) and (7)-(11) drop countries one-by-one and columns (6) and (12) jointly drops all countries. Columns (1)-(6) use specification in column (1) in Table 3; columns (7)-(12) use specifications in column (5) in Table 3.

Table B.14: Alternative Measure of Ancestry Composition Normalized by Country Size

	I(N.Supp>0)				I(N.Cust>0)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ancestry-to-Country Population Ratio 2010 (%)	0.399**	0.400**	0.297**	0.454**	0.334**	0.335**	0.233***	0.360**
	(0.189)	(0.190)	(0.120)	(0.213)	(0.137)	(0.138)	(0.078)	(0.142)
First-stage F stat	52.8	52.7	41.4	67.3	52.8	52.7	41.4	67.3
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	612495	612495	459150	612300	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we use alternative measure of ancestry composition defined as the ancestry-to-country population ratio in 2010. Therefore, the measure normalizes the origin-destination-level number of ancestry with the origin country size. Columns (1)-(4) use the specification in columns (1)-(4) in Table 3; columns (5)-(8) use the specifications in columns (5)-(8) in Table 3.

Table B.15: Impact of Ancestry Composition on Competitor Relationships

	I(N.Link>0)			
	(1)	(2)	(3)	(4)
Log Ancestry 2010	0.029*	0.029*	0.029	0.031*
	(0.017)	(0.017)	(0.019)	(0.019)
First-stage F stat	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-
Agricultural Similarity	-	-	✓	-
Origin x State FE	-	-	-	✓
Observations	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we use a linkage dummy based on competitor relationships as a dependent variable. Columns (1)-(4) use the specifications in columns (1)-(4) in Table 3, where only the dependent variable is replaced with the linkage dummy based on competitor relationships.

Table B.16: Non-manufacturing - Service-related Sectors

	I(N.Supp>0)				I(N.Cust>0)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Ancestry 2010	0.199***	0.199***	0.212***	0.215***	0.174***	0.174***	0.182***	0.186***
	(0.023)	(0.023)	(0.028)	(0.024)	(0.022)	(0.022)	(0.024)	(0.022)
First-stage F stat	76.3	75.9	60.2	84.8	76.3	75.9	60.2	84.8
Principal Components	✓	✓	✓	✓	✓	✓	✓	✓
Destination x Continent FE	✓	✓	✓	✓	✓	✓	✓	✓
Origin x Census Division FE	✓	✓	✓	-	✓	✓	✓	-
3rd order poly in dist and lat	-	✓	✓	-	-	✓	✓	-
Agricultural Similarity	-	-	✓	-	-	-	✓	-
Origin x State FE	-	-	-	✓	-	-	-	✓
Observations	612495	612495	459150	612300	612495	612495	459150	612300

Notes. This table presents the coefficient estimates from the IV regressions of Equation (3.1), where we measure firms' US county location using trade-engaging non-manufacturing establishments classified as (i) transportation & public utilities (SIC 40-49), (ii) finance, insurance, real estate (SIC 60-67), and (iii) services (SIC 70-89). Columns (1)-(4) use the specifications in columns (1)-(4) in Table 3; columns (5)-(8) use the specifications in columns (5)-(8) in Table 3.

Table B.17: Heterogeneous Treatment Effect:
External Finance Dependence Indicator (Rajan and Zingales, 1998)

	External Finance Dependence	
	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)
Log Ancestry 2010	0.033	0.053**
	(0.022)	(0.022)
Log Ancestry 2010 × CC	0.208***	0.173***
	(0.040)	(0.040)
First-stage F stat	451.6	451.6
Destination FE	✓	✓
Origin FE	✓	✓
Principal Components	✓	✓
Observations	498615	498615

Notes. This table repeats the exercise in Table 7, where we use an alternative measure of credit constraints: the external finance dependence indicator from Rajan and Zingales (1998). To facilitate the interpretation of coefficients, the credit constraint variable is standardized so that the sample mean equals zero and the sample standard deviation equals one.

Table B.18: Heterogeneous Treatment Effect: Credit Constraints - Additional Controls: Distance and Ethnic Fractionalization

Panel A	Distance			
	100-PayDexMax		100-PayDexMin	
	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)	(3)	(4)
Log Ancestry 2010	-0.033	0.008	-0.035	0.009
	(0.038)	(0.044)	(0.047)	(0.048)
Log Ancestry 2010 × CC	0.179***	0.128***	0.157***	0.100***
	(0.043)	(0.035)	(0.044)	(0.029)
Log Ancestry 2010 × Log Distance	0.056**	0.046*	0.063**	0.053**
	(0.024)	(0.024)	(0.026)	(0.026)
First-stage F stat	224.1	224.1	183.6	183.6
Destination FE	✓	✓	✓	✓
Origin FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	592995	592995	592995	592995

Panel B	Ethnic Fractionalization			
	100-PayDexMax		100-PayDexMin	
	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(5)	(6)	(7)	(8)
Log Ancestry 2010	-0.035	0.001	-0.054	-0.009
	(0.040)	(0.041)	(0.044)	(0.042)
Log Ancestry 2010 × CC	0.198***	0.136***	0.178***	0.111***
	(0.050)	(0.045)	(0.052)	(0.038)
Log Ancestry 2010 × Fractionalization	0.107*	0.095**	0.144**	0.123**
	(0.055)	(0.047)	(0.057)	(0.049)
First-stage F stat	283.6	283.6	222.9	222.9
Destination FE	✓	✓	✓	✓
Origin FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	492642	492642	492642	492642

Notes. This table repeats the heterogeneous treatment effect results in Table 7, where we additionally include interactions of Log Ancestry 2010 and other initial characteristics. Specifically, Panel A additionally controls for the interaction with the origin-destination-level log distance, and Panel B controls for the interaction with the origin country-level ethnic fractionalization (Alesina et al., 2003).

Table B.19: Heterogeneous Treatment Effect: Credit Constraints - Additional Controls: Firm Size and Final Goods Sector Share

Panel A	Firm Size			
	100-PayDexMax		100-PayDexMin	
	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)	(3)	(4)
Log Ancestry 2010	0.023	0.063	0.031	0.072*
	(0.043)	(0.041)	(0.046)	(0.042)
Log Ancestry 2010 × CC	0.150***	0.155***	0.111**	0.100**
	(0.047)	(0.044)	(0.052)	(0.049)
Log Ancestry 2010 × Firm Size	-0.020	-0.001	-0.024	-0.009
	(0.018)	(0.016)	(0.022)	(0.021)
First-stage F stat	130.2	130.2	99.9	99.9
Destination FE	✓	✓	✓	✓
Origin FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	592995	592995	592995	592995

Panel B	Final Goods Sector Share			
	100-PayDexMax		100-PayDexMin	
	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(5)	(6)	(7)	(8)
Log Ancestry 2010	-0.081	0.020	-0.092	-0.010
	(0.130)	(0.110)	(0.115)	(0.107)
Log Ancestry 2010 × CC	0.188***	0.148**	0.158***	0.107***
	(0.060)	(0.058)	(0.056)	(0.041)
Log Ancestry 2010 × Final Goods Sector Share	0.311	0.117	0.366	0.232
	(0.352)	(0.294)	(0.302)	(0.266)
First-stage F stat	138.1	138.1	155.3	155.3
Destination FE	✓	✓	✓	✓
Origin FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	592995	592995	592995	592995

Notes. This table repeats the heterogeneous treatment effect results in Table 7, where we additionally include interactions of Log Ancestry 2010 and other initial characteristics. Specifically, Panel A additionally controls for the interaction with the county-level average firm size, and Panel B controls for the interaction with the county-level employment share of the final goods sector establishments.

Table B.20: Heterogeneous Treatment Effect at the Firm-Origin-Destination Level: Credit Constraints

		100-PayDexMax			100-PayDexMin		
	I(N.Link>0)	I(N.Link>0)	I(N.Supp>0)	I(N.Cust>0)	I(N.Link>0)	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Ancestry 2010	0.008***	-0.003	-0.007**	0.006*	-0.000	-0.004	0.008***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Log Ancestry 2010 × CC		0.088***	0.072***	0.038*	0.091***	0.072***	0.033*
		(0.018)	(0.018)	(0.020)	(0.023)	(0.021)	(0.019)
First-stage F stat	488.2	142.4	142.4	142.4	121.7	121.7	121.7
Destination x Sector FE	✓	✓	✓	✓	✓	✓	✓
Origin x Sector FE	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓	✓	✓	✓
Observations	4699364	4699364	4699364	4699364	4699364	4699364	4699364

Notes. This table presents results of running regressions at the firm-origin-destination level. The goal of the exercise is to check if higher ancestry composition from origin country o to destination county d increases the probability that firm f—who has a trade-engaging establishment in destination d—has supply chain linkages with origin o between 2011 and 2014. The regression specification is $Y_{f,o,d} = \delta_f + \delta_{o,s} + \delta_{d,s} + \beta A_{o,d}^{2010} + \gamma A_{o,d}^{2010} \times CC_{f,d} + X_{o,d}'\gamma + \varepsilon_{f,o,d}$, where $Y_{f,o,d}$ can take I(N.Supp>0), I(N.Cust>0), and I(N.Link>0). I(N.Supp>0) and I(N.Cust>0) are dummy variables that equal one if firm f, who has a trade-engaging establishment in destination county d, has at least one supplier (customer) headquartered in origin country o. I(N.Link>0) is a linkage dummy that incorporates both supplier and customer linkages as in Table B.4. Column (1) shows the average treatment effect. Columns (2)-(7) show the heterogeneous treatment effects based on the measure of initial credit constraints, $CC_{f,d}$. Specifically, $CC_{f,d}$ measures firm-destination-level credit constraints, which is obtained by calculating weighted average of "100-PayDex" across trade-engaging establishments within each firm-destination, weighted by establishments' employment. δ_f indicates firm fixed effects. $\delta_{o,s}$ ($\delta_{d,s}$) indicates origin (destination)-by-sector fixed effects, where sector is defined as a firm's primary SIC 4-digit industry. All regressions control for log distance, latitude difference, and origin-destination-level FDI dummy. Standard errors are clustered at the origin country-level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.21: Heterogeneous Treatment Effect at the Firm-Origin-Destination Level: Credit Constraints
- Additional Controls: Distance and Ethnic Fractionalization

	I(N.Link>0)			
	100-PayDexMax	100-PayDexMin	100-PayDexMax	100-PayDexMin
	(1)	(2)	(3)	(4)
Log Ancestry 2010	-0.014	-0.013	-0.015**	-0.014**
	(0.009)	(0.009)	(0.007)	(0.006)
Log Ancestry 2010 × CC	0.094***	0.098***	0.085***	0.087***
	(0.022)	(0.027)	(0.019)	(0.024)
Log Ancestry 2010 × Log Distance	0.006	0.007*		
	(0.004)	(0.004)		
Log Ancestry 2010 × Fractionalization			0.019**	0.021**
			(0.009)	(0.009)
First-stage F stat	102.0	116.5	95.9	85.5
Destination x Sector FE	✓	✓	✓	✓
Origin x Sector FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	4699364	4699364	3887432	3887432

Notes. This table repeats the heterogeneous treatment effect results in Table B.20, where we additionally include interactions of Log Ancestry 2010 and other initial characteristics. Specifically, Panel A additionally control for the interaction with the origin-destination-level log distance, and Panel B controls for the interaction with the origin country-level ethnic fractionalization (Alesina et al., 2003).

Table B.22: Heterogeneous Treatment Effect at the Firm-Origin-Destination Level: Credit Constraints
- Additional Controls: Firm Size and Final Goods Sector Dummy

	I(N.Link>0)			
	100-PayDexMax	100-PayDexMin	100-PayDexMax	100-PayDexMin
	(1)	(2)	(3)	(4)
Log Ancestry 2010	-0.025**	-0.014	0.035***	0.036***
	(0.012)	(0.009)	(0.011)	(0.013)
Log Ancestry 2010 × CC	0.073***	0.082***	0.101***	0.101***
	(0.019)	(0.023)	(0.020)	(0.023)
Log Ancestry 2010 × Firm Size	0.041*	0.025		
	(0.021)	(0.017)		
Log Ancestry 2010 × Final Goods Sector			-0.089***	-0.085***
			(0.026)	(0.031)
First-stage F stat	34.2	41.0	87.0	38.3
Destination x Sector FE	✓	✓	✓	✓
Origin x Sector FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	4699364	4699364	4698218	4698218

Notes. This table repeats the heterogeneous treatment effect results in Table B.20, where we additionally include interactions of Log Ancestry 2010 and other initial characteristics. Specifically, Panel A additionally control for the interaction with the firm size (measured by the log of initial firm employment), and Panel B controls for the interaction with the final goods sector dummy. Final goods sector dummy is defined at the county-firm level and takes value of one if a firm's average upstreamness across its establishments within a given county is below 2 as in Antras et al. (2012).

Table B.23: Heterogeneous Treatment Effect: Credit Constraints in Destination County and Financial Development of Origin Country

	100-PayDexMax		100-PayDexMin	
	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)	(3)	(4)
Log Ancestry 2010	0.033	0.063**	0.040	0.070***
	(0.030)	(0.027)	(0.032)	(0.024)
Log Ancestry 2010 × CC	0.188***	0.120***	0.165***	0.099***
	(0.050)	(0.042)	(0.050)	(0.033)
Log Ancestry 2010 × (-FD)	-0.007	-0.012	-0.016	-0.018*
	(0.012)	(0.009)	(0.012)	(0.010)
First-stage F stat	216.4	216.4	239.4	239.4
Destination FE	✓	✓	✓	✓
Origin FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	462232	462232	462232	462232

Notes. This table repeats the heterogeneous treatment effect results in Table 7, where we additionally include interactions of Log Ancestry 2010 and the country-level measure of "minus" financial development (i.e., financial weakness) from Beck (2002) and Manova (2013). To facilitate the interpretation of the coefficients, the credit constraint variables and the measure of financial development are standardized so that the sample mean equals zero and the sample standard deviation equals one.

Table B.24: Heterogeneous Treatment Effect: Credit Constraints in Destination County, Financial Development and Judicial Quality of Origin Country

	100-PayDexMax		100-PayDexMin	
	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)	(3)	(4)
Log Ancestry 2010	-0.019	0.003	-0.013	0.010
	(0.044)	(0.040)	(0.051)	(0.041)
Log Ancestry 2010 × CC	0.239***	0.175***	0.215***	0.151***
	(0.054)	(0.045)	(0.059)	(0.042)
Log Ancestry 2010 × (-FD)	0.110*	0.121**	0.106*	0.115**
	(0.059)	(0.057)	(0.064)	(0.058)
Log Ancestry 2010 × (-JQ)	-0.086*	-0.099*	-0.092	-0.102*
	(0.052)	(0.051)	(0.056)	(0.052)
First-stage F stat	49.1	49.1	43.4	43.4
Destination FE	✓	✓	✓	✓
Origin FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	404453	404453	404453	404453

Notes. This table repeats the heterogeneous treatment effect results in Table 7, where we additionally include interactions of Log Ancestry 2010 and other initial characteristics. Specifically, we additionally include both the interaction with the country-level measure of "minus" financial development (i.e., financial weakness) from Beck (2002) and Manova (2013) and the interaction with the country-level measure of "minus" judicial quality (i.e., judicial weakness) from Nunn (2007). To facilitate the interpretation of the coefficients, the credit constraint variables, the measure of financial development, and the measure of judicial quality are standardized so that the sample mean equals zero and the sample standard deviation equals one.

Table B.25: Heterogeneous Treatment Effect: Credit Constraints in Destination County and Judicial Quality of Origin Country

	100-PayDexMax		100-PayDexMin	
	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)	(3)	(4)
Log Ancestry 2010	0.029	0.055*	0.032	0.060**
	(0.029)	(0.028)	(0.032)	(0.027)
Log Ancestry 2010 × CC	0.201***	0.127***	0.180***	0.106***
	(0.049)	(0.049)	(0.050)	(0.039)
Log Ancestry 2010 × (-JQ)	-0.016	-0.021***	-0.023**	-0.026***
	(0.010)	(0.008)	(0.010)	(0.009)
First-stage F stat	290.4	290.4	322.8	322.8
Destination FE	✓	✓	✓	✓
Origin FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	437904	437904	437904	437904

Notes. This table repeats the heterogeneous treatment effect results in Table 7, where we additionally include interactions of Log Ancestry 2010 and the country-level measure of "minus" judicial quality (i.e., judicial weakness) from Nunn (2007). To facilitate the interpretation of the coefficients, the credit constraint variables and the measure of judicial quality are standardized so that the sample mean equals zero and the sample standard deviation equals one.

Table B.26: Heterogeneous Treatment Effect:
The Role of Financial Development Interacting with Credit Constraints

	100-PayDexMax		100-PayDexMin	
	I(N.Supp>0)	I(N.Cust>0)	I(N.Supp>0)	I(N.Cust>0)
	(1)	(2)	(3)	(4)
Log Ancestry 2010	0.006	0.032^{**}	-0.043^{*}	-0.004
	(0.014)	(0.013)	(0.024)	(0.022)
Log Ancestry 2010 $\times$ CC	0.256^{***}	0.199^{***}	0.349^{***}	0.264^{***}
	(0.029)	(0.023)	(0.052)	(0.037)
Log Ancestry 2010 $\times$ CC $\times$ (-FD)	0.133^{***}	0.153^{***}	0.279^{***}	0.251^{***}
	(0.032)	(0.033)	(0.063)	(0.047)
Log Ancestry 2010 $\times$ (-FD)	-0.075^{***}	-0.090^{***}	-0.164^{***}	-0.150^{***}
	(0.023)	(0.018)	(0.039)	(0.028)
CC $\times$ (-FD)	-0.003	-0.006	-0.008^{*}	-0.009^{*}
	(0.003)	(0.004)	(0.005)	(0.005)
First-stage F stat	11.0	11.0	51.0	51.0
Destination FE	✓	✓	✓	✓
Origin FE	✓	✓	✓	✓
Principal Components	✓	✓	✓	✓
Observations	462232	462232	462232	462232

Notes. This table presents the heterogeneous treatment effect results by including the interaction of Log Ancestry 2010 with credit constraint measures (destination level) and the measure of "minus" financial development (i.e., financial weakness) (origin level). We use the same specifications as that in Table 7, where we additionally include the triple interaction among Log Ancestry 2010, credit constraints, and financial weakness (Log Ancestry 2010 x CC x (-FD)); the interaction between Log Ancestry 2010 and financial weakness (Log Ancestry 2010 x (-FD)); and the interaction between credit constraints and financial weakness (CC x (-FD)). The measure of credit constraints (CC) is obtained as the average credit constraint of establishments within each destination county, and the measure of country-level financial development (FD) comes from Beck (2002) and Manova (2013). To facilitate the interpretation of the coefficients, the credit constraint variables and the measure of financial development are standardized so that the sample mean equals zero and the sample standard deviation equals one.