Breaking Invisible Barriers:
Does Fast Internet Improve Access to Input Markets?*

Banu Demir† Beata Javorcik‡ Piyush Panigrahi§

December 4, 2023

Please find the latest version here

Abstract

How do improvements in internet infrastructure affect supply chain organization and distribution of economic activity across space? We exploit an episode of massive investment in optical fibre networks across Turkish provinces to shed light on this question. Using rich micro-data on firm-to-firm transactions and information on expansion of optical fibre networks, we show that fast internet affects firms' supplier networks by facilitating doing business. Firms not only reallocate their purchases towards suppliers with better internet connectivity but also diversify their input sourcing patterns. We develop and estimate a tractable spatial equilibrium model with endogenous production networks under rational inattention. We find that the elasticity of manufacturing firm-to-firm trade with respect to internet connectivity is sizeable. Through the lens of the estimated model, we find that improvement in high-speed internet infrastructure leads to a 2% gain in the real income in the median Turkish province.

*We thank our discussants Andrew Bernard, Julien Martin, Andreas Moxnes, and Lin Tian for their insightful comments. We also thank conference and seminar participants at the EBRD, Warwick, Birmingham, Cambridge, Bank of Italy, Maastricht, KCL Trade and Development Workshop, STEG, ERWIT, BSE Summer Forum and CESIfo. We are also grateful to the EIS team at the Ministry of Industry and Technology of the Republic of Turkey for their support with the disclosure process, and Ookla for sharing with us their speed data for Turkey. All interpretations, errors and omissions are the authors' own.

†Oxford, Bilkent, and CEPR; e-mail:banu.demirpakel@economics.ox.ac.uk

‡Oxford, EBRD, and CEPR; email: beata.javorcik@economics.ox.ac.uk

§World Bank; e-mail:ppanigrahi@worldbank.org

1 Introduction

Recent technological progress and the COVID-19 pandemic have led to an accelerated adoption of digital technologies and boosted investment in ICT infrastructure to support remote work, business-to-business communications, e-commerce, online learning, and other digital activities. Internet access has fundamentally changed the way communication occurs in modern economies. It has made communication faster, more convenient, and more accessible, allowing people to connect with others in ways that were previously impossible. The availability of high-speed internet enabled remote face-to-face meetings among participants from multiple locations, exchanging large files, or simultaneous working on data located on shared drives or in the cloud. Teams consisting of people working for different companies and located in different places can seamlessly collaborate facilitating business-to-business transactions, exchange of information and troubleshooting, joint design, electronic billing, etc. Around the world, governments and multilateral development banks see digitalization as a powerful tool for promoting economic growth and development. In particular, the development of digital infrastructure, such as high-speed internet networks, is seen as a priority.

This paper studies empirically and theoretically the effects of fast and reliable internet access on input sourcing and hence economic growth across space. It is one of the very few studies using microdata on firm-to-firm transactions to assess the impact of ICT infrastructure improvements on firms' input sourcing patterns. It is also the first study assessing the welfare gains resulting from ICT infrastructure improvements in a general equilibrium setting and providing a comparison of its benefits vis a vis transportation infrastructure.

Our study is based on the empirical context of Turkey, a country that has invested heavily in high-speed internet infrastructure and significantly expanded its optical fibre cable network between 2012 and 2019. Leveraging extensive microdata on firm-to-firm transactions and the deployment of optical fibre cables across Turkish provinces, we find that firms exhibit a stronger propensity to acquire inputs from regions with superior internet connectivity. Further, firms diversify their input sourcing by engaging with a greater number of suppliers and distributing their sourcing more equitably among suppliers located in provinces with improved internet connectivity. We rationalize these findings with a spatial equilibrium model that incorporates rationally inattentive input sourcing by firms. We structurally estimate the model and find that internet connectivity not only reduces the cost of obtaining information about potential suppliers but also reduces costs of synchronous communication with suppliers. Using the estimated model, we find that gains from improvement in high-speed internet infrastructure in real income across Turkish provinces are smaller but of a comparable magnitude relative to those that would be accrued from improvements in transportation infrastructure.

Our analysis centers around the construction of an extensive microdata set sourced from Turkey, which comprises the following key components: i) province-level data encompassing fibre cable length, fixed broadband penetration, and mobile subscriber statistics for the period spanning 2012 to 2019, obtained from the Information and Communication Technologies Authority; ii) microdata covering a quasi-comprehensive set of firm-to-firm transactions, complemented by essential firm characteristics such as location (province) and industry of operation, along with balance sheet information including gross sales, employment figures, and wage information, obtained from the Ministry of Industry and Technology, spanning the same period of 2012 to 2019; and iii) the GIS database of information on the BOTAS natural gas and oil pipeline network, which serves as the backbone for the optical fibre cable network.

Armed with this database, the analysis proceeds in two parts. In the first part, we provide empirical evidence on the effect of the roll-out of optical fibre cables on firms' input sourcing at across Turkish provinces. The roll-out of optical fibre cables across Turkey over 2012-2019 was staggered across provinces albeit with a secular increase in the length of optical fibre cables rolled out in the median province. The roll-out of cables was accompanied by uniform adoption of high-speed internet by firms across all sectors. In this context, our empirical strategy aims to use the temporal variation in high-speed internet access across the cross-section of Turkish provinces to capture the effects on firms' input sourcing both across provinces and within provinces across individual suppliers. To do so, we postulate that to engage in high-quality communication and remote collaboration, both parties require access to high-speed internet. Using data on the length of optical fibre cables rolled out, we construct bilateral measures of internet connectivity across pairs of Turkish provinces for each year that captures this complementarity between internet speeds at both provinces.

In reduced form regressions, the identifying assumption is that the higher proximity to the optical fibre network does not affect input sourcing of firms at a destination province from an origin province except through their effect on better internet connectivity at both provinces. We assess the validity of this assumption in several ways. We report how point estimates are affected by the inclusion of a variety of controls across origin-destination province pairs, both time-varying and time-invariant, as well as by using alternative specifications of our bilateral measure. To address concerns that the roll-out of optical fibre cables was correlated with province pairs which were predisposed to trade more a priori, we employ estimation using instrumental variables.

Our instrumental variable is motivated by the context in which Turkey's massive expansion of optical fibre network occurred starting late 2011. On October 3, 2011, the government made a significant decision regarding Fibre Access Services, stipulating their exemption from regulations for a span of five years or until the proportion of fibre internet subscribers reached 25% of the fixed broadband subscriber base. The ensuing growth in optical fibre network was facilitated by the government's decision to grant private internet providers the authority to utilize optical fibre cables laid out by BOTAS, the local natural gas and oil distributor in Turkey, to connect to farther locations. Optical fibre cables are typically laid out along with oil and gas pipelines to enable close monitoring and detection of faults. The BOTAS optical fibre network was laid out for pipeline monitoring before the expansion was set in motion. Since the BOTAS network was not laid out to facilitate internet connectivity across provinces, we are able to exploit plausibly exogenous variation in the distance of individual districts in Turkey to the BOTAS network to construct our instrument for internet connectivity.

Using this design, we find that better internet connectivity has strong and significant positive effects on input trade by firms at a destination province from an origin province. That is, firms reallocate their input purchases towards provinces with better internet connectivity. Furthermore, they diversify their input sourcing by engaging with more suppliers and sourcing more equitably across suppliers, conditional on the reallocation across origins.

In the second part of the paper, we then shed light on the aggregate implications through a quantitative spatial equilibrium model featuring endogenous formation of input-output linkages between firms under rational inattention. We build on the theoretical framework developed in Oberfield (2018) and Panigrahi (2022a), and extend it in several dimensions to capture economic forces that are relevant in our context. In the model, firms' production processes consist of multiple input requirements, and firms select the most attractive suppliers for their production requirements. However, firms operate with imperfect information regarding the attractiveness of potential suppliers. A potential supplier is considered more attractive if it (a) possesses a higher match productivity with the firm, (b) has lower production costs, and (c) is geographically closer thus incurring lower trade costs. Nevertheless, acquiring information about these suppliers incurs costs. The firm is only privy to the distribution from which match productivities are drawn but can opt to invest attention resources to gather information about the marginal costs of potential suppliers, akin to the approach employed by Dasgupta and Mondria (2018). The presence of internet connectivity influences both match productivities and information costs among buyer-supplier pairs. Where both origin and destination provinces have relatively better internet connectivity, match productivities are stochastically dominant in a first-order sense, and information costs are lower. The former implies that firms reallocate their input purchases towards provinces with better internet connectivity, while the latter implies that firms diversify their supplier base more extensively in provinces with superior internet access.

To quantify these forces, we estimate the model parameters, and calibrate the model to the 2012 Turkish economy as a reference equilibrium. In particular, we estimate the elasticity of firm-to-firm trade with respect to internet connectivity using an approach that combines model-based maximum likelihood with the exclusion restrictions of our IV through control functions. We find that the estimated elasticity is positive, statistically significant and of a magnitude comparable to the elasticity with respect to travel time. We also report how point estimates of our structural parameters are affected by the inclusion of a variety of controls across origin-destination province pairs, both time-varying and time-invariant, as well as by using alternative specifications of bilateral internet connectivity.

Armed with the calibrated model, we proceed to explore general equilibrium counterfactuals. We find that better access to high-speed internet in 2019 relative to 2012 led to a 2% increase in the real income in the media Turkish province. To put this into perspective, we ask how real income change across Turkish provinces would change if the 2012 Turkish economy were subject to the same improvements in transportation infrastructure that occurred in Turkey during the period 2005-2010. We find that real income would increase by 10% in the median Turkish province from this exercise. While gains from internet infrastructure are smaller on average relative to that from transportation infrastructure, they are significant and exhibit higher dispersion across provinces. Not only is our paper the first study assessing the welfare gains resulting from ICT infrastructure improvements in a general equilibrium setting, but also the only one providing a comparison of its benefits vis a vis transportation infrastructure. It is also one of the very few studies using the quasi-universe of firm-to-firm transactions to assess the impact of ICT infrastructure improvements.

Our study contributes to several strands of literature. First, it is closely related to the literature examining the impact of internet on various economic outcomes (see Hjort and Tian (2021) for a comprehensive review), especially international trade. In a cross-country study, Freund and Weinhold (2004) find that increase in the number of web hosts led to export growth. Fernandes, Mattoo, Nguyen, and Schiffbauer (2019) show evidence that internet roll-out in China increased firm-level exports, while Malgouyres, Mayer, and Mazet-Sonilhac (2021) use the staggered roll-out of broadband internet in France to show its positive effect on firm-level imports. Hjort and Poulsen (2019) find evidence of a notable increase in direct exports when submarine internet cables reach Africa. Exploiting the roll-out of the global telegraph network, Juhász and Steinwender (2018) show evidence that improvements in ICT increased trade in intermediates whose specifications can easily be communicated at a distance. In another study, Akerman, Leuven, and Mogstad (2022) exploit the roll-out of broadband internet in Norway and show that availability of internet increases the sensitivity of trade to distance. Jiang (2023) finds that firms that adopt more advanced technology have both higher within-firm communication and larger geographic coverage. We contribute to this literature by highlighting the importance of bilateral communication costs, by showing that access to fast internet enhances firms' ability to access a wider variety of inputs also within national borders, and by relying on a quasi-universe of firm-to-firm transactions.

Second, our paper is also related to papers studying information frictions in trade models. Rauch and Trindade (2003) augment a conventional trade model with informational trade barriers to demonstrate how the Internet and other ICT technologies can enhance the compatibility of international trade partners. This, in turn, leads to a greater integration of labor markets. In a similar vein, Allen (2014) introduces information frictions into a trade model, positing that diverse producers undertake a costly, sequential search process to determine optimal markets for their goods. His findings suggest that information frictions play a crucial role and help explain observed trading patterns in empirical data. Dickstein and Morales (2018) provide evidence that exporters operate without complete information sets, with larger firms possessing superior knowledge of foreign market conditions. They observe that improved access to information leads to an increase in total exports, even as the number of exporters decreases. Meanwhile, Dasgupta and Mondria (2018) endogenize information within a trade model, revealing that information costs exert non-monotonic and asymmetric impacts on bilateral trade flows. Our model is closest in this literature to Dasgupta and Mondria (2018) in that we introduce information costs and rationally inattentive behavior in firms' input sourcing decisions but extend the formulation of information costs to allow for more flexible patterns of substitution across suppliers.

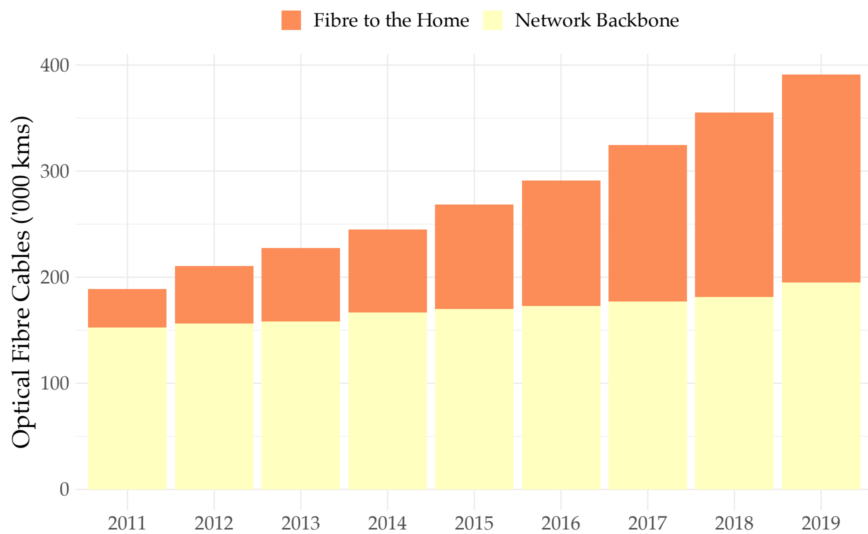
Third, our paper is related to the literature that studies the propagation of shocks in production networks (Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012); Baqaee and Farhi (2019a,b, 2020); Bigio and LaO (2020); Demir, Javorcik, Michalski, and Ors (2022)), and in particular ones that model endogenous formation of production networks (Chaney (2014); Lim (2018); Oberfield (2018); Huneeus (2020); Acemoglu and Azar (2020);

Demir, Fieler, Xu, and Yang (2021); Eaton, Kortum, and Kramarz (2022); Bernard, Dhyne, Magerman, Manova, and Moxnes (2022); Panigrahi (2022a,b); Miyauchi (2023); Arkolakis, Huneeus, and Miyauchi (2023)). Our contribution extends the framework outlined in Panigrahi (2022a) in two crucial ways. First, we introduce match-specific productivities among buyer-seller pairs that can stochastically vary contingent on the availability of high-speed internet access between the provinces where these firms are located. Second, we incorporate rationally inattentive behavior into firms' decisions regarding input sourcing. This means that firms endogenously determine the amount of information to acquire, with information acquisition costs being lower for province pairs with better internet connectivity.

Finally, our paper contributes the growing literature on market integration (Bernard, Moxnes, and Saito (2019, 2020); Donaldson (2015, 2018); Cristea (2011)) in quantitative spatial economics (Allen and Arkolakis (2014); Redding (2016); Caliendo, Parro, Rossi-Hansberg, and Sarte (2017a); Redding and Rossi-Hansberg (2017); Redding (2022); Cosar, Demir, Ghose, and Young (2021)). While this literature has focused on the impact of technology and transportation cost shocks on the spatial economy, we utilize the broader framework to assess the quantitative impact of reduction in communication costs due to high-speed internet access on the spatial distribution of economic activity. Cristea (2011), Bernard, Moxnes, and Saito (2019), and Bernard, Moxnes, and Saito (2020) show that lower transportation costs facilitate creation of new business relationships. Our contribution lies in demonstrating that access to fast internet connections can have similar effect on facilitating firm-to-firm interactions as lower transport costs. Exploiting two large-scale public investment programs within the same country – one in transport and the other in high-speed internet infrastructure, we are able to present a comparative welfare analysis of the two types of infrastructure improvements.

The remainder of the paper proceeds as follows. Section 2 describes the background of optical fibre cable rollout in Turkey and the data. Section 3 presents the empirical evidence on the effects of fibre cable rollout on firms' input sourcing patterns. Section 4 describes our model of input sourcing under rational inattention. Section 5 describes out estimation framework and presents results from structural estimation. Section 6 presents the theoretical framework that guides the welfare analysis. Section 7 presents our results from the quantitative assessment of the effects of optical fibre cable roll-out in Turkey. Section 8 concludes.

Figure 1: **Optical Fibre Cable Roll-out in Turkey**



Note: This figure depicts the roll-out of optical fibre cables and its breakdown between the backbone of the network and peripheral fibres laid to reach farther locations (fibre to the home) across Turkish provinces during the period 2012-2019. It is based on data obtained from the ICT Authority in Turkey. Over 2012-2019, the length of optical fibre cables rolled out increased by 85% with the network backbone increasing by 33% and that of cables rolled out to expand the network increasing by 375%

2 Background & Data

2.1 Rollout of Fibre Optic Infrastructure in Turkey

Prior to 2010, Turkey's internet infrastructure was extensive but had limited speed. On October 3, 2011, the government made a significant decision regarding Fibre Access Services, stipulating their exemption from regulations for a span of five years or until the proportion of fibre internet subscribers reached 25% of the fixed broadband subscriber base. Over the subsequent eight years, from 2011 to 2019, the fibre network expanded extensively, covering a remarkable distance of 390.8 thousand kilometers, equivalent to an impressive 0.48 kilometers per square kilometer of land area. This remarkable growth was facilitated by the government's decision to grant private internet providers the authority to utilize Botas, the local natural gas and oil distributor's fibre cables, which played a pivotal role in accelerating the rollout of fibre connectivity across the country. Figure 1 shows that the length of optical fibre cables rolled out in Turkey almost doubled from 2012 to 2019. The initial phase of Turkey's fibre internet roll-out was concentrated in "urban and economically attractive" areas but then expanded to universal coverage. It also shows that investment in fibre optic infrastructure was primarily directed towards rolling out fibre internet to farther locations and less towards the backbone of the fibre optic network. Figure 2 shows that, between 2012 and 2019, not only has the number of subscribers (both households and firms) to fibre internet lines increased fivefold, but their share in all fixed broadband connections has also been steadily rising. As of 2020, the fibre internet lines accounted for 23.9% of all fixed broadband connections in Turkey, converging to the OECD average of 30.6%.

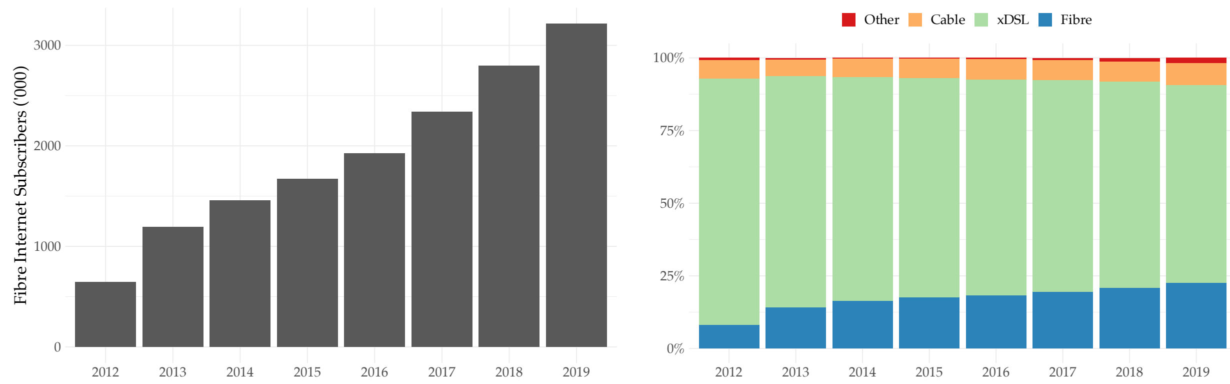
2.2 Data

This subsection provides a brief overview of the main datasets used in the analysis. We combine data from multiple sources. Broadly, they fall into four categories: (a) data relating to internet availability and usage across Turkish provinces, (b) data on firm-to-firm transactions and other firm-level outcomes, (c) GIS data on pipeline and road networks, and (d) administrative economic data.

High-Speed Internet Data

For the period 2012-2019, we have (i) data on ICT usage from a survey of approximately 10,000 firms that includes all firms with more than 250 employees and a representative sample of smaller firms, (ii) province-level information on length of optical fibre cables rolled out as well as number of fixed broadband, mobile and cable TV subscribers. The survey data on ICT usage comes from an annual firm-level survey conducted by the Turkish Statistical Institute (TUIK) each year since 2005. We use the survey data on ICT usage to assess firm-level adoption of high-speed internet which is defined as connections with internet speeds exceeding 100Mbps. The data on fibre cable roll-out and broadband subscribers is obtained from the Information and Communication Technologies Authority (ICTA) in Turkey. We use data on cable roll-out to construct measures of fibre internet connectivity across province-years and data on broadband subscriptions to construct measures of mobile internet connectivity and cable TV connectivity across province-years. For a shorter period 2016-2019, we also obtain data on upload and download speeds across finer locations in Turkey which we aggregate to the province-level to concord with the rest of the analysis. This data is obtained from Ookla, a private internet speed testing firm.

Figure 2: **Fibre Internet Subscribers**



Note: The left panel depicts the evolution of the number of fibre internet subscribers in Turkey during the period 2012-2019. The right panel shows the breakdown of fixed broadband connections into fibre, xDSL, Cable TV and others. Over 2012-2019, not only did the number of subscribers increase five-fold, but the share of broadband subscriptions due to fibre internet also increased.

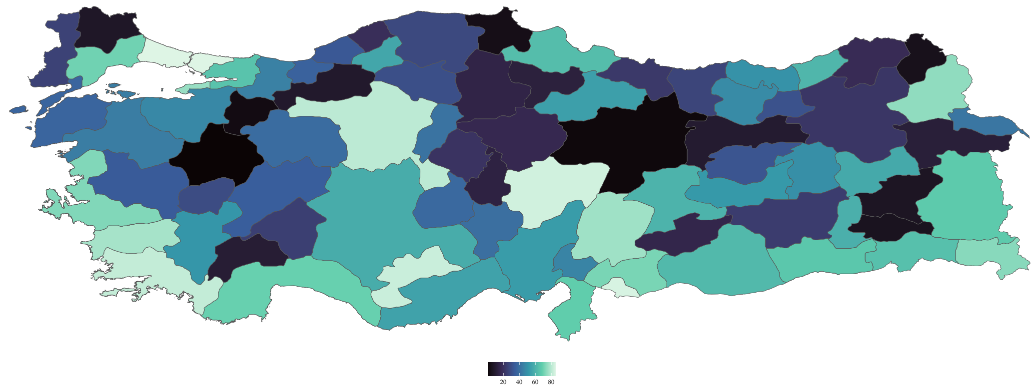
Firm-to-Firm Trade & Firms' Balance Sheet Data

We combine three data sets covering all formal firms in Turkey from 2012 through 2019. The Ministry of Industry and Technology (MoIT) in Turkey maintains all the data sets and uses the same firm identifier, allowing us to merge them. The data sets are as follows. First, the VAT data report the value of all domestic firm-to-firm trade that exceeds 5,000 Turkish liras (about US$840 in 2019) in a given year. Second, from the income statements, we use the yearly gross sales, employment, wages, and exports as well as imports of each firm. Third, from the firm registry, we extract each firm's province and 4-digit NACE code, the standard industry classification in the European Union. We restrict the analysis to firms in the manufacturing sector unless otherwise noted.

GIS Data

We digitize the map of BOTAS' oil and gas pipeline network in 2011 which also forms the backbone of the optical fibre network. Using this GIS data, we measure distance of Turkish districts to the pipeline network to construct our instruments for fibre connectivity. We also obtain the GIS database of the road network in Turkey to measure travel times between Turkish provinces in 2005 and 2010 (as in Cosar et al. (2021)). We use travel time data first as a control and then to construct identified shocks in travel time when comparing welfare effects of high-speed internet versus transportation infrastructure.

Figure 3: **Change in Optical Fibre Length**



Note: This figure depicts the spatial variation in optical fibre cable roll-out across Turkish provinces during the period 2012-2019. The median Turkish province saw a 68% increase in optical fibre roll-out, with a maximum of 177% for Istanbul and a minimum of 31% for Kutahya.

Administrative Economic Data

We obtain data on economic outcomes such as sectoral GDP, population, employment rate, urbanization rate etc from TUIK both at the province and district level. Turkey has 81 provinces which are further subdivided into 973 districts. We use this data as controls for spatial and temporal variation in province and district characteristics.

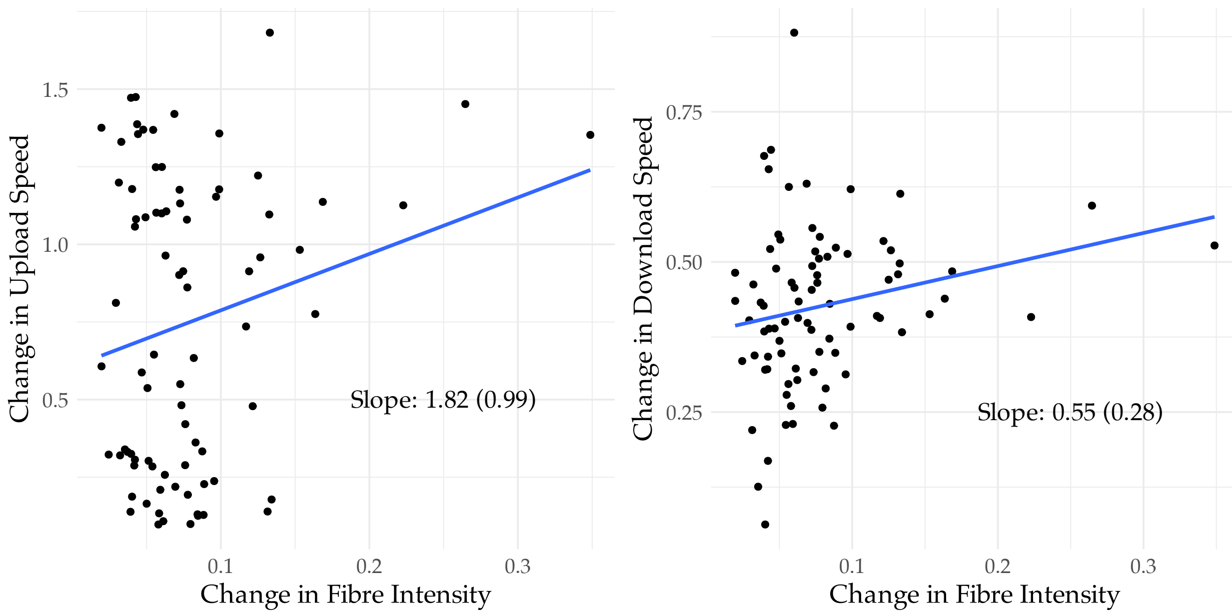
2.3 Spatial Variation in Fibre Internet Rollout

To capture spatial variation in roll-out of fibre optic infrastructure across Turkish provinces and years, we construct a measure of fibre intensity for each province and year. In particular, we measure fibre intensity as the length of fibre optic cables rolled out in a province normalized by its land area. Formally, for a province d in year t, fibre intensity I_{dt} is calculated as:

$$I_{dt} = \ln\left(1 + \frac{L_{dt}}{A_d}\right),$$

where L_{dt} denotes the length of optical fibre cables (in kms) rolled out in province d in year t and A_d is the surface area of the province (in km^2). Figure 3 depicts the change in fibre intensity across Turkish provinces between 2012 and 2019.

Figure 4: **Correlation of Fibre Intensity with Upload and Download Speeds**



Note: The left panel depicts how change in measured fibre intensity correlates with change in reported upload speeds during the period 2016-2019. The right panel depicts correlation of fibre intensity with reported download speeds. Data on upload and download speeds across Turkish provinces during 2016-2019 was obtained from Ookla.

We assess whether more fibre optic cables translate to better internet connectivity in Turkish provinces. To do so, we obtain upload and download speeds reported across provinces for years 2016-2019. Figure 4 shows that change in our measure of fibre intensity correlates positively with both upload and download speeds.

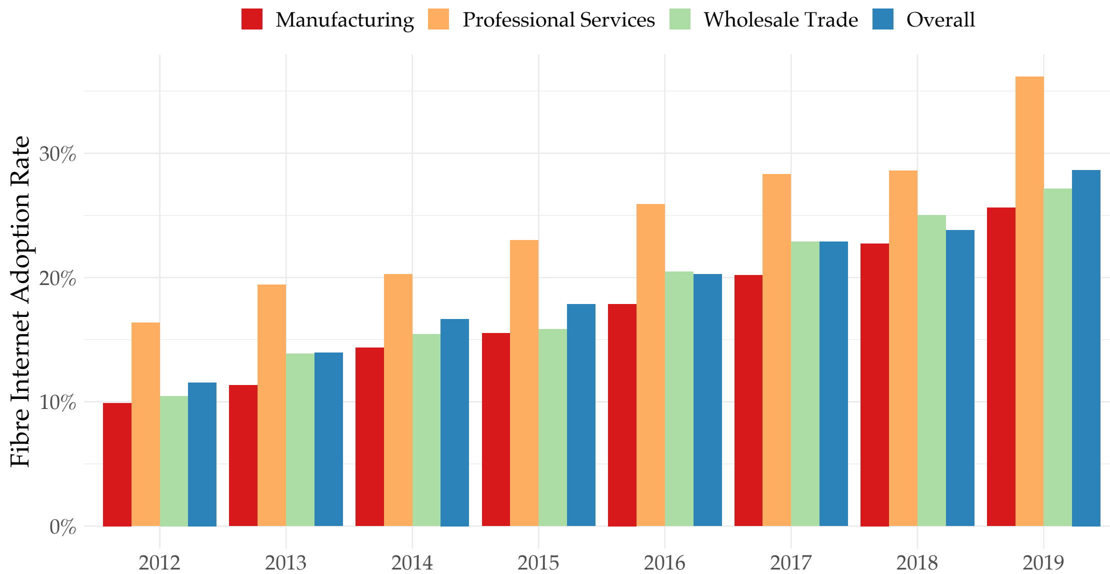
2.4 Fibre Internet Adoption by Firms

The roll-out went hand-in-hand with high take-up by firms. Figure 5 shows that the share of firms with high-speed internet followed the trend in roll-out and increased drastically since 2011.[1] The fraction of firms with high-speed internet increased from zero in 2011 to about 30% by the end of 2019.

To investigate how strongly the adoption responded to the roll-out of fibre internet, we estimate the following equation for the 2012-2019 period:

[1]We consider a firms to have access to high-speed internet if it reports internet speed over 100Mbps in the ICT Survey.

Figure 5: **Firms' Adoption of High-Speed Internet in Turkey**



Note: This figure depicts the evolution of firms' adoption of high-speed internet across broad sectors (manufacturing, professional services and wholesale trade) and for the overall economy. During 2012-2019, firms in all sectors increasingly adopt high-speed internet. The fraction of firms with high-speed internet increased from zero in 2011 to about 30% by the end of 2019.

$$\text{Adoption of High-Speed Internet}_{ft} = \gamma I_{dt} + \alpha_d + \alpha_{st} + \text{Size}_{ft} + e_{ft} \tag{1}$$

where the dependent variable takes on the value one if firm f adopts high-speed internet in year t, and zero otherwise. The specification controls for firm size in terms of employment and includes province as well as sector-year fixed effects.[2] In Table A1, column (1), the parameter of interest, which captures the adoption rate by firms, is positive and statistically significant, suggesting that firms adopted high-speed internet as it became available in their provinces. The magnitude of the estimate implies that a 1 percent increase in fibre cable length to area is associated with a 0.017 percentage-point increase in the share of firms with high-speed internet access, which corresponds to a 8.7 percent increase relative to the mean value the dependent variable over the sample period.

In Table A1 column (2), we estimate the effect of the availability of high-speed internet on firm-level adoption by sector.The results imply that firm-level adoption of high-speed internet as a response to its availability does not vary significantly across sectors.

[2] Sectors are aggregated from 2-digit NACE industries as listed in the second column of Table A1.

3 Empirical Evidence

This section uses the datasets described above to estimate the effects of high-speed internet access on the spatial distribution of firms' intermediate input purchases.

3.1 Empirical Strategy

Turkey witnessed a massive investment in fibre optic infrastructure during 2012-2019. This resulted in staggered roll-out of optical fibre cables across Turkish provinces. In this context, our aim is to exploit both temporal and cross-sectional variation in fibre connectivity across province pairs to capture the consequences of high-speed internet access on input sourcing patterns and hence supply chain organization across Turkish provinces.

Fibre-optic internet is often considered superior to other broadband technologies like DSL (Digital Subscriber Line) or cable internet for several reasons. First, fibre internet offers faster download and upload speeds compared to many other broadband technologies. It can provide symmetrical speeds, meaning the upload speed is as fast as the download speed. This makes fibre-optic internet ideal for bandwidth-intensive activities like high-definition streaming and large file transfers. Second, fibre-optic connections generally have lower latency compared to some other broadband options. Low latency is crucial for real-time applications like video conferencing, as it reduces delays and lag in communication. Third, fibre-optic cables are less susceptible to interference from electrical and radio frequency sources, making them more reliable than some other broadband technologies, especially in areas with high levels of electromagnetic interference. Finally, fibre internet offers a more consistent and reliable speed experience compared to other technologies like DSL or cable. While internet speed with DSL or cable may be influenced by factors like distance from the provider's equipment or network congestion, fibre provides a more stable and predictable performance.

For synchronous communication to work effectively, whether it's in the form of a video call, voice call, online chat, or any other real-time interaction both parties typically require a reasonably good internet connection. Synchronous communication involves the exchange of data in real-time. A stable and sufficiently fast internet connection is necessary to transmit this data smoothly. A good internet connection ensures that messages or signals are sent and received promptly. With a poor connection, there can be delays in sending and receiving messages, audio dropouts, video stuttering, which can disrupt the flow of conversation and make it less effective. A weak or unreliable connection can also lead to disconnects, dropped calls, or interrupted chats, which can be frustrating and disruptive to the communication process. Different forms of synchronous communication have varying bandwidth requirements. For example, video calls and high-quality voice calls require more bandwidth than text-based chat. If one party has limited bandwidth, it may struggle to handle the data requirements of the chosen communication method, leading to a suboptimal experience. Therefore, for smooth and effective synchronous communication, both parties should ideally have good internet connections. Based on this rationale, we construct a bilateral measure of connectivity across Turkish province pairs that aims to capture this strong complementarity of internet speed in both provinces. In particular, we proxy for fibre connectivity between provinces as the minimum of fibre intensity between both provinces, that is,

$$\text{Fibre Connectivity } I_{od,t} = \min\left\{I_{o,t}, I_{d,t}\right\}$$

To estimate the effect of fibre connectivity on firms' input sourcing strategy across province pairs over time, we estimate the following baseline specification:

$$\ln y_{od,t}(b) = \beta \ln \text{Fibre Connectivity}_{od,t} + \alpha_{bt} + \alpha_{ot} + \alpha_{od} + \alpha' X_{od,t} + \epsilon_{od,t}(b) \tag{2}$$

where o indexes origin provinces, b indexes buyer firms, d indexes destination province where b is located and t indexes years. In our baseline specification, we regress origin-buyer-year level outcomes $y_{od,t}(b)$ on our bilateral measure of fibre connectivity, buyer-year fixed effects, origin-year fixed effects, and origin-destination fixed effects. To address concerns about auto-correlated error terms for the same province pair over time, we cluster standard errors at the origin and destination level. After reporting the reduced-form estimation results, we then estimate IV point estimates using an instrument for fibre connectivity. Subsequently, to account to the substantial number of zero observations at the origin-buyer-year levels, we estimate the specification using PPML and also report point estimates obtained via a control function approach to address endogeneity concerns. The identifying assumption in specification (2) is that length of fibre cable rolled out affect buyer-origin level outcomes relative to other buyer-origin pairs only through their effect of internet connectivity from fibre internet. To assess this assumption, as we discuss in detail below, we also report a number of additional robustness checks as part of the reduced-form and IV estimation.

3.2 Reduced-Form Estimation

Reallocation across Origins

We begin by estimating the effect of differences in fibre connectivity on firms' input sourcing strategy across provinces. To do so, for each buyer firm and year, we compute the share of its material costs that is due to purchases from suppliers from each origin province in that particular year. Using data on firm-to-firm transactions that we obtain from the Ministry of Industry and Technology in Turkey, the cost share of a firm b located in province d due to purchases from suppliers at province o in year t is calculated as:

$$\text{Cost Share}_{od,t}(b) = \frac{\sum_{s \in o} \text{Purchases}_{od,t}(s, b)}{\sum_{o'} \sum_{s' \in o'} \text{Purchases}_{o'd,t}(s', b)},$$

where $\text{Purchases}_{od,t}(s, b)$ denotes the value of transactions made by firm b with supplier s located in province o.

We estimate specification (2) with these cost shares on the left-hand side. Table 1 Panel A, column (1) presents the reduced-form results. We find that the effect of fibre connectivity on cost share from origin is positive and statistically significant at the 1% level. This implies that Turkish firms sourced inputs relatively more from provinces for which fibre connectivity improved.

Diversification across Suppliers within Origins

To further investigate the effect of fibre connectivity on input sourcing, we explore the effect of fibre connectivity on input sourcing across suppliers within an origin province. We consider three outcomes. First, we estimate the effect of fibre connectivity on the number of suppliers at province o that firm b in province d sources from in year t. Table 1 Panel A, column (2) reports the reduced-form results. We find that fibre connectivity between provinces o and d has a positive and statistically significant positive effect on the number of suppliers that b engages with in o.

Second, we estimate the effect of fibre connectivity on the concentration of input purchases across suppliers within an origin province. To do so, for each buyer firm and year, we compute the Herfindahl-Hirschmann index of its share of material costs that is due to purchases from suppliers from each origin province in that particular year. In particular, the cost share HHI of a firm b located in province d due to purchases from suppliers at province o in year t is calculated as:

$$\text{Cost Share HHI}_{od,t}(b) = \sum_{s \in o} \left(\frac{\text{Purchases}_{od,t}(s,b)}{\sum_{s' \in o} \text{Purchases}_{od,t}(s',b)} \right)^2 .$$

We estimate specification (2) with these cost share HHIs on the left-hand side. Table 1 Panel A, column (3) presents the reduced-form results. We find that the effect of fibre connectivity on cost share HHI from origin is negative and statistically significant at the 1% level.

Finally, we estimate the effect of fibre connectivity on the number of new connections that a firm makes at an origin province in a given year relative to the year before. Table 1 Panel A, column (4) presents the reduced-form results. We find the effect to be positive and statistically significant at the 1% level.

These results imply that not only do Turkish firms reallocate their input purchases towards provinces with better internet connectivity but they also diversify their input sourcing strategy conditional on reallocation. They source from more suppliers and do so more equitably across suppliers.

Robustness Checks

Table 1 Panel A documents strong effects of fibre connectivity on firms' spatial distribution of input purchases captured by cost share from origin, number of suppliers, cost share HHI across provinces. One potential concern is that other correlated factors that simultaneously affect fibre connectivity and firms input sourcing might affect our estimates. We subject our estimates in Table 1 to a battery of robustness checks.

With the inclusion of buyer-year, origin-year, and origin-destination fixed effects, any potential threat to identification would come from time-varying factors pertaining to an origin-destination pair. In Table A2, we include (a) absolute difference in GDP per capita at origin and destination to account for the fact that similar provinces are likely to trade more and (b) mobile connectivity between origin and destination provinces to account for other forms of connectivity that might make it more likely for origin and destination provinces to trade more. The results show that our baseline estimates are robust to omission of such variables.

In Table A3, we use an alternative measure of fibre connectivity computed as the negative absolute difference of fibre intensity at origin and destination provinces. We find that the coefficients are still statistically significant and of the same sign as the baseline estimates. In Table A4 we conduct a placebo test where we replace our fibre connec-

Table 1: **Reallocation and Diversification in Input Sourcing**

Dependent Variable:	Cost Share (1)	No. Suppliers (2)	Cost Share HHI (3)	New Connections (4)
Panel A: OLS				
Fibre Connectivity	0.510***	0.325***	-0.107***	0.0632***
	(0.0549)	(0.0297)	(0.0136)	(0.0196)
Panel B: 2SLS				
Fibre Connectivity	0.498***	0.309***	-0.102***	0.0629***
	(0.0577)	(0.0310)	(0.0139)	(0.0197)
KP test stat.	31.9	31.9	31.9	31.9
Fixed Effects:				
Buyer×Year	✓	✓	✓	✓
Origin×Year	✓	✓	✓	✓
Origin×Destination	✓	✓	✓	✓
Observations	2,230,473	2,230,473	2,230,473	2,230,473

Note: Each observation pertains to a buyer firm, an origin province and a year. All variables are in natural logarithms. Cost Share is the fraction of purchases of a buyer from the origin province. No. Suppliers is the number of suppliers of the buyer firm located in a given origin province. Cost Share HHI is the Herfindahl-Hirschman Index of cost shares of suppliers of a buyer, which are located in a given origin province. New Connections is the number of new suppliers relative to the year before. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

tivity variable with the minimum value of the ratio of cable TV subscribers per capita at origin and destination. While both require similar infrastructure, cable TV does not provide direct benefits to firms. To the extent that our results reflect the adoption of high-speed internet by businesses rather than improvements in regional infrastructure over the sample period, the extent of cable TV subscription should not affect trade-related outcomes. The results confirm our conjecture. In Table A5, we also include purchases from non-manufacturing suppliers and find that our results still hold in sign and statistical significance.

In Table A6, we look into the source of variation in our bilateral connectivity measure. Our measure is dominated by the trade partner whose fibre connectivity is of lower quality than the other. Therefore, we expect that improvements in fibre connectivity of a trade partner have an effect on our bilateral connectivity measure when the other trade partner's connectivity measure is in the higher quartiles of the respective distribution. The results presented in Table A6 confirm our predictions.

In addition to the analysis presented here, we present further robustness results in the

IV estimation that follows.

3.3 IV Estimation

One potential concern is that FTTH predominantly reached economically attractive locations or locations that are predisposed to trade more a priori. To address this, we employ instrumental variables estimation.

Construction of IV

Our choice of instrumental variable is driven by the context in which Turkey witnessed a substantial expansion of its optical fibre network, commencing in late 2011. Specifically, on October 3, 2011, the government made a significant decision concerning Fibre Access Services. This decision entailed exempting these services from regulatory constraints for a period of five years or until the proportion of fibre internet subscribers reached 25% of the fixed broadband subscriber base, whichever came first. This decision paved the way for the subsequent growth in the optical fibre network, as it empowered private internet providers to utilize optical fibre cables laid out by BOTAS, the local natural gas and oil distributor in Turkey, to establish connections to more distant locations.

Optical fibre cables are commonly installed alongside oil and gas pipelines to serve as reliable communication infrastructure. These fibre cables are used by operators to detect potential leaks or damages, identify security breaches or any attempts to tamper with the pipeline infrastructure and effectively manage the substantial volumes of data generated by sensors and monitoring equipment distributed along the length of the pipeline. The BOTAS optical fibre network was laid out for before the expansion of the optical fibre network set in motion. Since the BOTAS network was not laid out to facilitate internet connectivity across provinces, we are able to exploit plausibly exogenous variation in the distance of individual districts in Turkey to the BOTAS network to construct our instrument for internet connectivity.

The 81 provinces of Turkey are further subdivided into 973 districts. For each district m, we calculate the minimum distance from the district center to the BOTAS pipeline network, Z_m. We then compute the average distance for each province to the pipeline network as the weighted average distance of its districts with weight pertaining to each district as its share of the province's total population in 2011. In particular, the weighted distance of province o to the pipeline network is calculated as:

$$Z_o = \sum_{m \in o} \frac{\text{Population}_{m,2011}}{\text{Population}_{o,2011}} \times Z_m.$$

We show that these distance measures across provinces are not correlated with a series of initial province characteristics, such as GDP, area, population, manufacturing share of GDP, urbanization rate, and employment rate, pertaining to the year 2011. Figure A5 shows that the corresponding estimates are not significantly different from zero.

To capture the notion of strong complementarity in internet speed for synchronous communication, we construct the IV for fibre connectivity for each province pair as:

$$Z_{od} = \max \{Z_o, Z_d\}.$$

Since this bilateral measure is time-invariant, we use its interactions with year dummies for years 2012 through 2019 in our first-stage regression. One would expect a negative correlation between this measure and our bilateral fibre connectivity measure, and also that the magnitude of the correlation to increase over time. The results presented in Figure A6 are consistent with our prior: relative to the year 2012, distance to the BOTAS pipeline network is negatively related to high-speed internet connectivity across all years with the effect increasing over time. The interaction terms are jointly significant, with an F-statistic above 500.

IV Estimation Results

Table 1, Panel B presents IV point estimates of the effect of fibre connectivity on cost share, number of suppliers, cost share HHI, and number of new connections. In the IV estimation, the effect of fibre connectivity is positive and statistically significant for firms' cost share from origin, number of suppliers and new connections while negative and statistically significant for cost share HHI among suppliers at origin. Reassuringly, the IV point estimates are close to the OLS estimates.

4 A Model of Input Sourcing under Rational Inattention

With these empirical results in hand, we now lay out a spatial equilibrium framework that features endogenous formation of input-output linkages between spatially distant firms, whose main objectives are twofold. First, estimation of the model would shine light on the quantitative importance of high-speed internet access in reducing trade frictions. Second, the estimated model would enable quantification of gains in real income that can be attributed to increased internet access between the years 2012 and 2019. In what follows, we describe a model of trade between provinces in Turkey. Each province consists of a positive measure of firms. Firms produce using local labor and intermediate inputs sourced from suppliers potentially spread across multiple provinces. Firms operate under imperfect information when deciding on their input sourcing but can expend attention to acquire more information about potential suppliers. Attention is costly, therefore, firms are rationally inattentive, i.e., they balance their expected payoff against the cost of information. Households are endowed with one unit of labor and supply inelastically to firms located in their province of residence. Trade between provinces is subject to iceberg trade costs, that is, a firm producing at o needs to ship τ_{od} units of a good for one unit of good to arrive at d.

Throughout the paper, a firm is indexed by s when it is a *seller* of intermediate inputs or goods for final consumption and by b when it is a *buyer* of intermediate inputs. A province is indexed by o when it is the *origin* of a trade flow and typically where firm s is located. Similarly, it is indexed by d when it is the *destination* of a trade flow and typically where firm b is located. The set of all locations is denoted by $\mathcal{J}$. The set of all firms is denoted by $\mathcal{M}$ and the subset located at o is denoted by $\mathcal{M}_o$. The number of elements in these sets are denoted as $M = |\mathcal{M}|$ and $M_o = |\mathcal{M}_o|$. All proofs are relegated to the appendix.

4.1 Technology and Market Structure

Firms' production processes involve combining labor and accomplishing a set of tasks by sourcing intermediate inputs from other firms as in Eaton et al. (2022) and Panigrahi (2022a). In particular, the production function for any firm b at province d is defined over labor and a discrete number of tasks (indexed by $k \in \mathcal{K}_d(b) \equiv \{1, \cdots, K_d(b)\}$) as:

$$y_d(b) = z_d(b) \left(\frac{l_d(b)}{1-\alpha_d} \right)^{1-\alpha_d} \left(\frac{\prod_{k \in \mathcal{K}_d(b)} m_d(b,k)^{1/K_d(b)}}{\alpha_d} \right)^{\alpha_d},$$

$$m_d(b,k) = \sum_{s \in \mathcal{M}} m_{od}(s,b,k),$$

where $y_d(b)$ is the output of firm b, $l_d(b)$ is the amount of labor input used, $m_d(b,k)$ is the quantity of materials utilized to accomplish task k, $z_d(b)$ is the idiosyncratic Hicks-neutral productivity with which firm b produces, α_d is the materials share of costs, and $K_d(b)$ is the number of tasks in the production function of firm b. For accomplishing any task, the outputs of potential suppliers are perfectly substitutable. The elasticity of substitution across suppliers is different from the elasticity of substitution across tasks. The former is determined by how likely it is to find a more suitable supplier for each task while the latter is assumed to be unity for simplicity.

4.2 Supplier Choice

Having described the production function in terms of tasks, we now turn to supplier choice for each of the tasks. For each of its tasks, the supplier that firm b selects depends on three factors: (a) marginal cost of the supplier $c_o(s)$, (b) iceberg trade cost of shipping goods from the origin to the destination τ_{od} and (c) match-specific productivity, $a_{od}(s,b,k)$. Firms operate under imperfect information about marginal costs of potential suppliers and the distribution of match-specific productivities with potential suppliers from a particular origin province, when making supplier choice decisions. Each firm b has some prior knowledge about the available options, given by a probability measure ν. For each task k, the conditional probability of choosing supplier s is denoted by $\pi_{od}(s,b,k)$. Given the conditional probabilities $\boldsymbol{\pi}_d(b,k)$ and the prior ν, the unconditional probabilities are defined as $\bar{\boldsymbol{\pi}}_d(b,k) = \mathbb{E}_\nu\left[\boldsymbol{\pi}_d(b,k)\right]$. The unconditional probabilities are also the prior or expected probability of choosing the supplier. The problem of the rationally inattentive firm b is to choose the vector of conditional probabilities $\boldsymbol{\pi}_d(b,k) \equiv \{\pi_{od}(s,b,k)\}$ balancing the expected cost of production against the cost of information. We specify information costs $\psi_d(b,k)$ in terms of mutual Bregman information as follows:

$$\psi_d(b,k) = \Omega\left(\bar{\boldsymbol{\pi}}_d(b,k)\right) - \mathbb{E}_\nu\left[\Omega\left(\boldsymbol{\pi}_d(b,k)\right)\right],$$

where $\Omega(\cdot)$ denotes generalized entropy as proposed by Fosgerau, Melo, de Palma, and Shum (2020). The higher the difference in generalized entropies associated with the conditional and unconditional vector of probabilities, the higher is the information cost. The generalized entropy, based on Bregman (1967), for a vector of conditional probabilities $\boldsymbol{\pi}_d(b,k)$ is defined as $\Omega\left(\boldsymbol{\pi}_d(b,k)\right) = -\boldsymbol{\pi}_d(b,k) \cdot \log \boldsymbol{S}\left(\boldsymbol{\pi}_d(b,k)\right)$ where $\boldsymbol{S}(\cdot)$ is a mapping that satisfies $S\left(\boldsymbol{\pi}_d(b,k), s\right) = 0$ iff $\pi_{od}(s,b,k) = 0$. The generalized entropy for the vector of unconditional probabilities is similarly defined.[3]

With this specification of information cost, the input sourcing strategy of firm b is a

[3]This formulation of information cost is a generalization of the one based on Shannon (1948) entropy in Sims (2003) and Matejka and McKay (2015) where the mapping $\boldsymbol{S}(\cdot)$ is set to be the identify function. In recent work, Dasgupta and Mondria (2018) and Bertoli, Moraga, and Guichard (2020) utilize the framework developed in Matejka and McKay (2015) to explain trade and migration patterns respectively.

solution to the following problem:

$$\log p_d(b,k) = \max_{\pi_d(b,k)} \left\{ -\mathbb{E}_\nu \left[\sum_{s \in \mathcal{M}} \pi_{od}(s,b,k) \ln \left(\frac{c_o(s)\tau_{od}}{a_{od}(s,b,k)} \right) \right] - \psi_d(b,k) \right\}, \tag{3}$$

$$\text{subject to} \quad \pi_{od}(s,b,k) \geq 0 \text{ for all } s \in \mathcal{M},$$

$$\sum_{s \in \mathcal{M}} \pi_{od}(s,b,k) = 1,$$

where $p_d(b,k)$ is the effective price of task k for firm b. The solution to the above problem are posterior choice probabilities after the firm takes into account the additional costly information it acquired. Taking wage w_d and effective prices $\{p_d(b,k) : k \in \mathcal{K}_d(b)\}$ as given, the firm's unit cost function is given by:

$$c_d\,(b) = \frac{w_d^{1-\alpha_d} \left(\prod_{k \in \mathcal{K}_d(b)} p_d(b,k)^{1/K_d(b)} \right)^{\alpha_d}}{z_d(b)}. \tag{4}$$

4.3 High-Speed Internet Access and Input Sourcing

We assume that quality of internet access available to both the origin and destination affects both information costs and match-specific productivities via two functional form assumptions. First, we assume that match-specific productivities are more likely to be higher when both the origin and destination have higher access to fibre internet. To model this relationship, we assume that these productivities are drawn independently for all potential suppliers involved in each of the tasks within a firm's production functions. This process follows the Fréchet distribution, which we define in the following assumption.

Assumption 1. *Match-specific productivities of firms at province o for tasks of firms at d are drawn independently according to the following Fréchet distribution:*

$$F_{a,od}(a) = \exp\left(-\phi_{od} a^{-\zeta}\right),$$

$$\phi_{od} = \exp\left(\bar{\gamma} + \gamma \ln I_{od}\right)$$

In our assumption, we specify that the scale parameter of the Fréchet distribution is unique to each origin-destination pair and is assumed to have a constant elasticity concerning bilateral fibre connectivity. This means that the match-specific productivities are drawn from distributions that are stochastically better if both the origin and destination have higher access to fibre internet. Simply put, the better the quality of internet access between the origin and destination provinces, the more likely it is for the match-specific productivities to be higher. When making their information acquisition decision, firms attempt to obtain more information about the scale parameter ϕ_{od}. The exact realization of match-specific productivities would still remain unknown.

Second, we assume that costs of acquiring information about potential suppliers are lower when both the origin and destination have higher access to fibre internet. Further, since iceberg trade costs and scale parameter for match-specific productivities are common across suppliers at an origin province, obtaining information about any supplier also reduces additional information required about other suppliers in the same province. To capture this, we specify the mapping $\boldsymbol{S}(\cdot)$ as follows.

Assumption 2. *The mapping $\boldsymbol{S}(\boldsymbol{\pi}_d(b,k)) \equiv \{S(\boldsymbol{\pi}_d(b,k),s) : s \in \mathcal{M}\}$ is defined as:*

$$S(\boldsymbol{\pi}_d(b,k),s) = \lambda_{od} \ln \pi_{od}(s,b,k) + (1-\lambda_{od}) \ln \left(\sum_{s' \in \mathcal{M}_o} \pi_{od}(s',b,k) \right),$$

$$\lambda_{od} = \frac{1}{1+\eta \ln I_{od}},$$

where λ_{od} is the information cost parameter for firms at d when selecting suppliers at o.

In our assumption, we specify that the information cost parameter is unique to each origin-destination pair and is assumed to be decreasing in bilateral fibre connectivity with a constant elasticity. In other words, the better the quality of internet access between the origin and destination provinces, the lower is the cost incurred by buyers at the destination province to obtain information about potential suppliers at the origin province.

4.4 Closing the Model

Household Preferences

Households are modeled analogously with tasks in their utility function. They select the most cost-effective suppliers for each task under imperfect information similar to firms sourcing inputs. Each household supplies one unit of labor inelastically to local firms and receives labor income.

Equilibrium Definition

Let $\sigma \equiv \{\boldsymbol{z}, \boldsymbol{\tau}, \boldsymbol{\phi}, \mathcal{K}, \boldsymbol{a}\}$ denote the aggregate state of the economy. Here $\boldsymbol{z}$ denotes the vector of idiosyncratic productivities of firms, $\boldsymbol{\tau}$ denotes the vector of trade costs across all pairs of provinces, $\boldsymbol{\phi}$ denotes the vector of match scale parameters across all pairs of provinces, $\mathcal{K}$ denotes the sets of tasks of all firms and households, and $\boldsymbol{a}$ denotes the vector of all match-specific productivities. All of these objects are exogenous. An equilibrium in this economy is an allocation and prices such that (a) households and firms select suppliers for tasks; (b) firms set prices for other firms and households under marginal cost pricing; (c) households maximize utility; (d) firms minimize costs; and (e) market clears for each firm's goods and for labor at each province. This completes description of the economic environment in the model.

Moving ahead, the aggregate state can be divided into two parts. The first comprises of firms' productivities, sets of tasks of all firms and households, trade costs, and match scale parameters $\sigma_0 \equiv \{\boldsymbol{z}, \boldsymbol{\tau}, \boldsymbol{\phi}, \mathcal{K}\}$. The second part comprises of match-specific productivities; this is denoted by $\sigma_1 \equiv \{\boldsymbol{a}\}$. While σ_0 narrows down the set of exogenous objects that the firm or household endogenously choose to obtain infomation about, σ_1 consists of objects that still remain unknown at network formation.

5 Taking Model to Data

To map the model to micro-data on firm-to-firm sales for estimation, we proceed in two steps. First, we utilize the recursive representation of network formation between firms to cast it as a quasi- dynamic programming problem and show that the model delivers closed-form characterization of conditional choice probabilities in this quasi- dynamic discrete choice setting. Second, we describe how these conditional choice probabilities coupled with multiple discrete choice across tasks lead to a high-dimensional nested logit model of supplier choice which can be estimated feasibly by exploiting special features of the corresponding likelihood specification. The resulting estimation framework is scalable and circumvents computational difficulties pervasive in estimation of network formation models with large numbers of firms.

In what follows, terms with no superscript denote true values, those with $(\cdot)^D$ denote data counterparts, and those with superscript $(\cdot)^*$ denote corresponding estimates. For example, $\pi_{od}(s, b)$ denotes true values of conditional choice probabilities, $\pi^D_{od}(s, b)$ denotes observed cost shares, and $\pi^*_{od}(s, b)$ denotes estimates of conditional choice probabilities.

5.1 Conditional Choice Probabilities & Firm-to-Firm Trade

We begin by casting network formation between firms as a quasi-dynamic programming problem. In particular, rewriting equation (4), we find that marginal cost of any firm b

admits the following recursive representation.

$$\ln c_d(b) = -\ln z_d(b) + (1-\alpha_d)\ln w_d \\ + \frac{\alpha_d}{K_d(b)} \sum_{k\in\mathcal{K}_d(b)} \max_{\pi_d(b,k)} \left\{ -\mathbb{E}_\nu \left[\sum_{s\in\mathcal{M}} \pi_{od}(s,b,k) \ln\left(\frac{c_o(s)\tau_{od}}{a_{od}(s,b,k)} \right) \right] - \psi_d(b,k) \right\} \tag{5}$$

This representation is akin to a setting with dynamic discrete choice. The estimands in this estimation problem are trade costs $\{\tau_{od} : (o,d) \in \mathcal{J}^2\}$, match scale $\{\phi_{od} : (o,d) \in \mathcal{J}^2\}$, and information costs $\{\lambda_{od} : (o,d) \in \mathcal{J}^2\}$ which are exogenous, and firms' marginal costs $\{c_o(s) : s \in \mathcal{M}\}$ which are endogenously determined, unobserved in the data and run into millions. We utilize the conditional choice probability approach to estimate the model following Hotz and Miller (1993). In this context, conditional choice probabilities are the probabilities with which any given supplier s is chosen for any one of the buyer b's tasks conditional on its marginal cost being $c_o(s)$.

We proceed to show next that the model delivers closed-form predictions for these probabilities. We derive expressions for conditional choice probabilities and hence predictions for firm-to-firm trade. Recall from equation (3) that firms choose suppliers for tasks based on suppliers' marginal costs, trade costs faced by them, and match-specific productivities associated with the task under consideration. While trade costs $\boldsymbol{\tau}$ and match scale $\boldsymbol{\phi}$ constitute σ_0, match-specific productivities are unknown and suppliers' marginal costs $c_o(s)$ are determined endogenously. We therefore characterize conditional choice probabilities for supplier choice, i.e., probabilities for choice of supplier conditional on its marginal cost but in expectation over match-specific productivities that are yet to be realized. Since match-specific productivities are independent and identically distributed across tasks for a given firm b, the probability of firm s getting selected for any one of the tasks by firm b is the same as for any other. Let $\pi_{od}(s,b)$ denote the probability with which firm b selects firm s for any one of its tasks. The following proposition provides expressions for conditional choice probabilities $\pi_{od}(s,b)$.

Proposition 1. *For any realization of σ_0, conditional on firm s's marginal cost being $c_o(s)$, the probability with which any firm b located in d selects firm s located in o for any given task is*

$$\pi_{od}(s,b) = \pi_{od}(s \mid o,b) \times \pi_{od}(\bullet,b), \\ \pi_{od}(s \mid o,b) = \frac{\bar{\pi}_{od}(s \mid o,b) c_o(s)^{-\zeta/\lambda_{od}}}{\sum_{s'\in\mathcal{M}_o} \bar{\pi}_{od}(s' \mid o,b) c_o(s')^{-\zeta/\lambda_{od}}}, \tag{6}$$

$$\pi_{od}(\bullet,b)=\frac{\bar{\pi}_{od}(\bullet,b)\tau_{od}^{-\zeta}\phi_{od}\left(\sum_{s'\in\mathcal{M}_o}\bar{\pi}_{od}(s'\mid o,b)c_o(s')^{-\zeta/\lambda_{od}}\right)^{\lambda_{od}}}{\sum_{o'}\bar{\pi}_{o'd}(\bullet,b)\tau_{o'd}^{-\zeta}\phi_{o'd}\left(\sum_{s'\in\mathcal{M}_{o'}}\bar{\pi}_{o'd}(s'\mid o',b)c_{o'}(s')^{-\zeta/\lambda_{o'd}}\right)^{\lambda_{o'd}}},\tag{7}$$

where $\bar{\pi}_{od}(s\mid o,b)=\frac{\bar{\pi}_{od}(s,b)}{\sum_{s'\in\mathcal{M}_o}\bar{\pi}_{od}(s',b)}$ *and* $\bar{\pi}_{od}(\bullet,b)=\sum_{s\in\mathcal{M}_o}\bar{\pi}_{od}(s,b)$.

The posterior choice probabilities have a structure similar to a nested logit, except that they are adjusted by the prior probabilities $\bar{\boldsymbol{\pi}}_d(b)\equiv\{\bar{\pi}_{od}(s,b):s\in\mathcal{M}\}$. Equation (6) provides the probability of choosing a supplier within a province while equation (7) provides the probability of choosing a province. Equation (6) shows that firms with lower marginal costs are more likely to get selected. The sensitivity of the probability with respect to marginal costs is decreasing in λ_{od}. That is, $|\frac{\partial\ln\pi_{od}(s|o,b)}{\partial\ln c_o(s)}|=\frac{\zeta}{\lambda_{od}}$. With better fibre connectivity, information costs λ_{od} are lower and the choice of supplier is more sensitive to marginal cost. Equation (7) shows that province pairs with better fibre connectivity are more likely to trade more intensively. The elasticity of the probability with respect to match scale is positive and equal to unity, i.e., $\frac{\partial\ln\pi_{od}(\bullet,b)}{\partial\ln\phi_{od}}=1$.

The tractable expressions for firm-to-firm trade in Proposition 1 give rise to transparent estimating equations for the model, to which we turn next.

5.2 A Nested Logit Model of Supplier Choice

We reformulate the economic model developed so far as a nested logit model of supplier choice for tasks of each of the firms and estimate it semi-parametrically with seller fixed effects and origin-destination fixed effects. Origin-destination fixed effects correspond to a structural gravity specification for estimating trade frictions. Trade frictions are then estimated by projecting origin-destination fixed effects on observables that capture geographic proximity such as distance and borders etc as well as on observables that capture high-speed internet access.

Making use of Proposition 1, the estimating equation can be expressed as a nested logit function:

$$\mathbb{E}\left[\pi_{od}^{D}(s,b)\right]=\pi_{od}(s,b)\tag{8}$$

Formally, the estimation problem is as follows:

$$\Delta^*=\arg\max_{\Delta}\frac{1}{M}\sum_{b\in\mathcal{M}}\ln f_{\mathrm{NL}}\left(\mathbb{D}\mid\Delta\right),\tag{9}$$

$$f_{\mathrm{NL}}\left(\mathbb{D} \mid \Delta\right) \propto \prod_{s \in \mathcal{M}} \left(\pi^{*}_{od}(s,b)\right)^{\pi^{D}_{od}(s,b)},$$

where

$$\Delta \equiv \left\{\left\{c_o(s)^{-\zeta} : s \in \mathcal{M}\right\}, \left\{\lambda_{od}, \phi_{od}, \tau_{od}^{-\zeta} : (o,d) \in \mathcal{J}^2\right\}\right\} \text{ and}$$
$$\mathbb{D} \equiv \left\{\pi^{D}_{od}(s,b) : (s,b) \in \mathcal{M}^2\right\}.$$

The above specification with fixed effects however presents a problem of perfect multi-collinearity in regressors. Note that dummy variables associated with $\{c_o(s)^{-\zeta} : s \in \mathcal{M}_o\}$ and $\left\{\phi_{od}\tau_{od}^{-\zeta} : d \in \mathcal{J}\right\}$ are collinear for all such provinces o. Hence, we make the following normalizations so that these fixed effects are identified up to scale. For all $s \in \mathcal{M}_o, o \in \mathcal{J}$, let $c_o(s) = c_o\widetilde{c}_o(s)$ be such that

$$\left(\sum_{s \in \mathcal{M}_o} \bar{\pi}_{od}(s \mid o,b)\widetilde{c}_o(s)^{-\zeta/\lambda_{od}}\right)^{-\lambda_{od}/\zeta} = 1 \quad \forall (o,d) \in \mathcal{J}^2$$

Fixed Effects and Structural Gravity

The first order conditions implied by the likelihood maximization problem in equation (9) can be solved to obtain closed-form estimators for fixed effects as described in the proposition below.

Proposition 2. *The estimates from equation (9) are given by:*

$$\left(\frac{\bar{\pi}_{od}(s \mid o,b)\widetilde{c}_o(s)^{-\zeta/\lambda_{od}}}{\sum_{s' \in \mathcal{M}_o} \bar{\pi}_{od}(s' \mid o,b)\widetilde{c}_o(s')^{-\zeta/\lambda_{od}}}\right)^{*} = \frac{1}{M_d}\sum_{b \in \mathcal{M}_d} \pi^{D}_{od}(s \mid o,b) \quad \forall s \in \mathcal{M}, d \in \mathcal{J}, \tag{10}$$

$$\left(\frac{\bar{\pi}_{od}(\bullet,b)c_o^{-\zeta}\phi_{od}\tau_{od}^{-\zeta}}{\sum_{o'} \bar{\pi}_{od}(\bullet,b)c_{o'}^{-\zeta}\phi_{od}\tau_{o'd}^{-\zeta}}\right)^{*} = \frac{1}{M_d}\sum_{b \in \mathcal{M}_d} \pi^{D}_{od}\left(\bullet,b\right) \quad \forall (o,d) \in \mathcal{J}^2 \tag{11}$$

where $\pi^{D}_{od}\left(\bullet,b\right) \equiv \sum_{s \in \mathcal{M}_o} \pi^{D}_{od}\left(s,b\right)$ *and* $\pi^{D}_{od}(s \mid o,b) = \frac{\pi^{D}_{od}(s,b)}{\pi^{D}_{od}(\bullet,b)}$.

Trade Frictions and Conditional Supplier Choice Probabilities

Trade frictions can now be estimated by projecting bilateral origin-destination fixed effects (from equation (11)) on bilateral observables such as travel time, similar to gravity regressions, in additional to variables that capture high-speed internet access with the following estimating equation:

$$\mathbb{E}\left[\left(\frac{\bar{\pi}_{od}(\bullet,-)c_o^{-\zeta}\phi_{od}\tau_{od}^{-\zeta}}{\sum_{o'}\bar{\pi}_{od}(\bullet,-)c_{o'}^{-\zeta}\phi_{o'd}\tau_{o'd}^{-\zeta}}\right)^*\right]=\frac{\exp\left(\ln\left(c_o^{-\zeta}\right)+\boldsymbol{\omega}_{od}'\boldsymbol{\gamma}+\boldsymbol{X}_{od}'\boldsymbol{\beta}+\tilde{\pi}_{od}\right)}{\sum_{o'}\exp\left(\ln\left(c_{o'}^{-\zeta}\right)+\boldsymbol{\omega}_{od}'\boldsymbol{\gamma}+\boldsymbol{X}_{o'd}'\boldsymbol{\beta}+\tilde{\pi}_{od}\right)}. \tag{12}$$

This delivers estimates of origin fixed effects $\left(c_o^{-\zeta}\right)^*$ and trade frictions arising from geography $\left(\tau_{od}^{-\zeta}\right)^* = \exp\left(\boldsymbol{X}_{od}'\boldsymbol{\beta}^*\right)$ and internet access $(\phi_{od})^* = \exp\left(\boldsymbol{\omega}_{od}'\boldsymbol{\gamma}^*\right)$. Estimates of conditional supplier choice probabilities are then estimated as per the following specification:

$$\mathbb{E}\left[\bar{\pi}_{od}(s \mid o,-)\widetilde{c}_o(s)^{-\zeta/\lambda_{od}}\right]=\exp\left(\ln\left(c_o^{-\zeta}\right)+\eta\ln I_{od}\times\ln\left(c_o^{-\zeta}\right)+\tilde{\pi}_{od}(s)\right) \tag{13}$$

Formally, the estimates of conditional choice probabilities are then given by

$$\pi_{od}^*(s,b)=\bar{\pi}_{od}^*(s \mid o,b)\left(\left(\widetilde{c}_o(s)^{-\zeta}\right)^*\right)^{1/\lambda_{od}^*}\cdot\pi_{od}^*(\bullet,b), \tag{14}$$

$$\pi_{od}^*(\bullet,b)=\frac{\bar{\pi}_{od}^*(\bullet,b)\left(c_o^{-\zeta}\right)^*\phi_{od}^*\left(\tau_{od}^{-\zeta}\right)^*}{\sum_{o'\in\mathcal{J}}\bar{\pi}_{od}^*(\bullet,b)\left(c_{o'}^{-\zeta}\right)^*\phi_{o'd}^*\left(\tau_{o'd}^{-\zeta}\right)^*}. \tag{15}$$

5.3 Estimation Results

We present estimation results from specification (12) in Table 2. In Columns (1) and (4) we include origin-destination fixed effects whereas in column (3) we control for bilateral travel time between provinces. The estimated elasticity with respect to fibre connectivity varies between 0.47-0.63 depending on the specification and sample in columns (1), (3) and (4). The estimated elasticity with respect to travel time is 1.6 which is larger than the estimated elasticity with respect to internet access. To address the potential endogeneity of the fibre internet investments, we take the control function approach and re-run our preferred specification in Columns (2) and (5) while controlling for the estimated residuals obtained from the first stage regression of our IV. While the estimates of the elasticity with respect to internet access are smaller than the baseline estimates both when including and excluding non-manufacturing suppliers, they are still sizeable and statistically significant.

Table A10 subjects our estimates in Table 2 column (1) to three robustness checks. In Column (1), we add bilateral travel time interacted with yearly dummies to the specification. In Column (2), we include absolute difference in GDP per capita at origin and destination to account for the fact that similar provinces are likely to trade more. In column (3), we include mobile connectivity between origin and destination provinces to account for other forms of connectivity that might make it more likely for origin and destination provinces to trade more. The results presented in columns (1)-(3) show that while the elasticity with respect to fibre connectivity is slightly reduced, it is still economically and statistically significant. In column (4), we use an alternative measure of fibre connectivity computed as the negative absolute difference of fibre intensity at origin and destination provinces. We find that while the elasticity is around half that of our original measure, it is still positive and statistically significant. In column (5) we conduct a placebo test where we replace our fibre connectivity variable with the minimum value of the ratio of cable TV subscribers per capita at origin and destination. While both require similar infrastructure, cable TV does not provide direct benefits to firms. To the extent that our results reflect the adoption of high-speed internet by businesses rather than improvements in regional infrastructure over the sample period, the extent of cable TV subscription should not affect trade-related outcomes. The results confirm our conjecture.

5.4 Connecting to Empirical Evidence

To enable the model to make predictions on the number of suppliers at an origin province among buyers at destination, we assume that firms' number of tasks are independently distributed according to a zero-truncated Poisson distribution as we define in the following assumption.

Assumption 3. *The number of tasks are drawn independently according to the following zero-truncated Poisson distribution:*

$$\mathbb{P}\left(K_d(b) = K\right) = \frac{e^{-\kappa}\kappa^K}{\left(1 - e^{-\kappa}\right)K!}.$$

We are now ready to state our proposition that links model predictions to empirical results.

Proposition 3. *As internet connectivity increases at an origin province o and a destination province d:*

1. *a firm located in d is likely to source a relatively larger proportion of its input requirements from suppliers located in o.*

2. *a firm located in d is likely to source from a larger number of suppliers located in o.*

Table 2: **Model-Consistent Gravity Regressions**

Dependent Variable:	Average Cost Share				
	Manufacturing		Overall		
	(1) PPML	(2) PPML:CF	(3) PPML	(4) PPML	(5) PPML:CF
Fibre Connectivity	0.629***	0.421***	0.470***	0.594***	0.277***
	(0.143)	(0.107)	(0.115)	(0.118)	(0.0672)
Travel Time			-1.550***		
			(0.0112)		
Fixed Effects:					
Origin×Year	✓	✓	✓	✓	✓
Destination×Year	✓	✓	✓	✓	✓
Origin×Destination	✓	✓		✓	✓
R^2	0.986	0.986	0.842	0.994	0.996
Observations	45504	45504	51840	51840	51840

Note: Each observation pertains to an origin province, a destination province, and a year. The dependent variable in columns (1)-(2) is constructed using firms' purchases from manufacturing suppliers. In columns (3)-(5), we use purchases from both manufacturing and non-manufacturing suppliers. To address endogeneity concerns, columns (2) and (5) are estimated via a control function approach by including predicted residuals obtained from the first-stage regression. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

3. *the concentration of input sourcing across suppliers at o among firms at d declines.*

6 Aggregation

For aggregation and counterfactual analysis, we adopt the large economy model due to Al-Najjar (2004) which is characterized by a sequence of finite but increasingly large economies $\{\mathcal{E}_t : t \in \mathbb{N}\}$ that progressively discretizes the unit continuum. We now proceed to characterize wages $\boldsymbol{w}(\sigma)$ in equilibrium in the limiting economy, i.e., $\lim_{t\to\infty} \mathcal{E}_t$.

6.1 Relative Wages in Trade Equilibrium

To define relative wages in trade equilibrium, we begin by characterizing sourcing probabilities, that is, the probability with which any buyer sources inputs from province o for any one its tasks. Conditional choice probabilities of supplier choice naturally aggregate to sourcing probabilities, that is, sourcing probabilities can be obtained as the sum of conditional choice probabilities associated with all the suppliers located at o. Conditional choice probabilities from Proposition 1 lead to the next proposition. This proposition characterizes sourcing probabilities across origins by firm b, denoted by $\pi_{od}(\bullet, b)$.

Proposition 4. *For any realization of σ_0, the probability with which any firm b located in d selects a supplier from o for any given task is*

$$\pi_{od}(\bullet, b) = \frac{\bar{\pi}_{od}(\bullet, -)\mu_o w_o^{-\zeta(1-\alpha_o)} \overline{z_o^{\zeta}} \mathbb{E}\left[\Gamma\left(1 - \frac{\alpha_o}{K_o(\cdot)}\right)^{K_o(\cdot)}\right] A_o^{\alpha_o} \phi_{od} \tau_{od}^{-\zeta}}{A_d}, \tag{16}$$

where $\boldsymbol{A} \equiv \{A_d : d \in \mathcal{J}\}$ is the unique positive solution to the following fixed point problem:

$$A_d = \sum_{o \in \mathcal{J}} \bar{\pi}_{od}(\bullet, -)\phi_{od}\tau_{od}^{-\zeta} \overline{z_o^{\zeta}} \mu_o w_o^{-\zeta(1-\alpha_o)} \mathbb{E}\left[\Gamma\left(1 - \frac{\alpha_o}{K_o(\cdot)}\right)^{K_o(\cdot)}\right] A_o^{\alpha_o}, \tag{17}$$

where μ_o denotes the proportion of firms at o and $\overline{z_o^{\zeta}} = \mathbb{E}\left[z_o(s)^{\zeta}\right]$.

These sourcing probabilities are independent of the identity of the buyer at the destination and therefore can be written as $\pi_{od}^0(\bullet, -)$. In the limiting economy, the average sourcing share across all buyers in the limiting economy coincides with the expected value given by equation (16). This however does not mean that the sourcing shares across individual buyers are identical either in the finite economy or the limiting economy. Buyers at a destination may very well differ in their sourcing shares whether in the finite economy or the limiting economy. Formally, the law of large numbers implies that in the limiting economy,

$$\frac{1}{M_d} \sum_{b \in \mathcal{M}_d} \pi_{od}(\bullet, b) \xrightarrow{t \to \infty} \pi_{od}(\bullet, -). \tag{18}$$

We now turn to characterizing relative wages in the trade equilibrium in the limiting economy. The following proposition shows that relative wages in the limiting economy can be obtained as a solution to the system of equations (19).

Proposition 5. *For any realization of $\sigma \equiv \{\sigma_0, \sigma_1\}$, $\boldsymbol{w} \equiv \{w_d : d \in \mathcal{J}\}$ solves the following system of equations:*

$$\frac{w_o L_o}{1 - \alpha_o} = \sum_{d \in \mathcal{J}} \pi_{od}(\bullet, -) \frac{w_d L_d}{1 - \alpha_d}. \tag{19}$$

Further, for any σ and σ' such that $\sigma_0 = \sigma_0'$ and $\sigma_1 \neq \sigma_1'$:

$$\boldsymbol{w} = \boldsymbol{w}'. \tag{20}$$

The above proposition also shows that, for any given realization of σ_0, relative wages are invariant across all networks realized for all values of σ_1. This concludes the characterization of equilibrium wages and brings us to the definition of the trade equilibrium below.

Definition 1. For any given σ_0, the trade equilibrium in the limiting economy is defined as the vector of wages $\boldsymbol{w}$ such that (a) market access at each province satisfies equation (17); (b) trade shares coincide with sourcing probabilities in equation (16) and (c) the market clearing condition in equation (19) holds.

7 Quantitative Analysis

7.1 Computation of Counterfactual Outcomes

We operationalize Propositions 4 and 5 for counterfactual analysis by expressing them in changes. The following definition states that and motivates the algorithm for evaluating counterfactual outcomes in response to shocks that derive from a change in the aggregate state σ_0 to σ_0'.

Definition 2. For any change in aggregate state σ_0 to σ_0', equilibrium change in wages $\widehat{\boldsymbol{w}} \equiv \{\widehat{w}_d : d \in \mathcal{J}\}$ and welfare $\widehat{\boldsymbol{V}} \equiv \left\{\widehat{V}_d : d \in \mathcal{J}\right\}$ are characterized the following system of equations for all realizations of σ_1 or σ_1':[4]

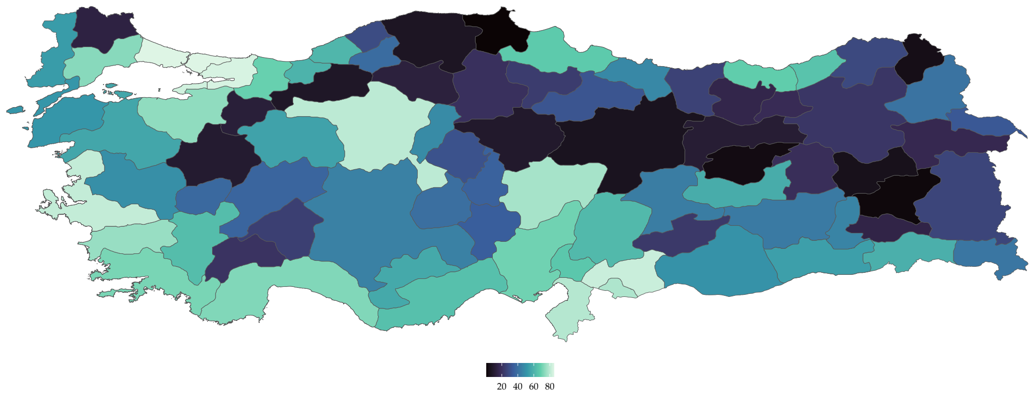
$$
\begin{aligned}
\widehat{A}_d &= \sum_o \pi_{od}\left(\bullet, -\right) \widehat{\tau_{od}}^{-\zeta} \widehat{\phi_{od}} \widehat{w}_o^{-\zeta(1-\alpha_o)} \widehat{A}_o^{\alpha_o} \\
\widehat{\pi_{od}\left(\bullet, -\right)} &= \frac{\widehat{\tau_{od}}^{-\zeta} \widehat{\phi_{od}} \widehat{w}_o^{-\zeta(1-\alpha_o)} \widehat{A}_o^{\alpha_o}}{\widehat{A}_d} \\
\frac{\widehat{w}_o w_o L_o}{1-\alpha_o} &= \sum_d \widehat{\pi_{od}\left(\bullet, -\right)} \pi_{od}\left(\bullet, -\right) \frac{\widehat{w}_d w_d L_d}{1-\alpha_d} \\
\widehat{V}_d &= \widehat{w}_d \widehat{A}_d^{1/\zeta}
\end{aligned}
$$

where $\widehat{\delta} \equiv \left\{\widehat{\delta}_{od} : (o,d) \in \mathcal{J}^2\right\}$ is function of shocks that capture the resultant effect of change from σ_0 to σ_0'.

With this definition of the equilibrium in changes in the limiting economy, aggregate and firm-level counterfactual outcomes in the limiting economy are computed in three steps. First, we evaluate aggregate outcomes in the limiting economy in the initial state.

[4] The expression for welfare changes is derived in Appendix E.1.

Figure 6: **Welfare Effects of High-Speed Internet Infrastructure**



Note: This figure depicts the spatial variation in welfare impact of high-speed internet roll-out across Turkish provinces during the period 2012-2019. The median Turkish province saw a 2% increase in welfare, with an interquartile range of 1% to 4%.

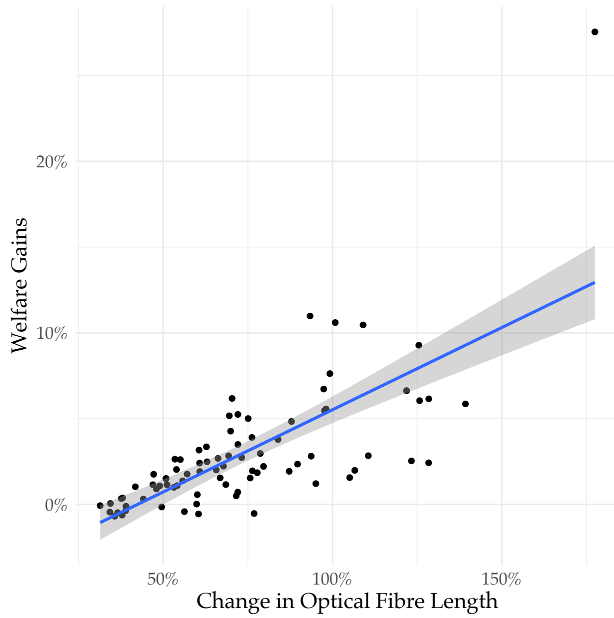
Second, we evaluate changes in aggregate outcomes when going from the initial state to the counterfactual state. This is done using a tâtonnement algorithm similar to Alvarez and Lucas (2007) and Dekle, Eaton, and Kortum (2008). Finally, we evaluate aggregate outcomes in the limiting economy in the counterfactual state. Details of the procedure are in Appendix E.2.

7.2 Welfare Effect of High-Speed Internet Infrastructure

Our objective is to conduct a quantitative assessment of the effect of optical fibre rollout in Turkey. To achieve this, we begin by selecting the Turkish economy in 2012 as our reference point. We start by constructing identified shocks to the economy using actual changes in fibre rollout between 2012 and 2019, and the elasticity of trade with respect to bilateral internet connectivity (measured as the minimum fibre intensity between province pairs) obtained from Column (4) in Table 2. Through the lens of our model, we evaluate how much real income would change across Turkish provinces if the 2012 economy was subject to these shocks ceteris paribus.

To determine the welfare gains resulting from these shocks, we use the procedure described in Section 7.1. We calibrate the trade elasticity ζ to 5 following the median of estimates in Head and Mayer (2014). Figure 6 provides a map of the welfare gains across the

Figure 7: **Welfare Gains vs Change in Optical Fibre Length**



Note: This figure plots welfare gains from high-speed internet infrastructure against percentage change in optical fibre cables rolled-out.

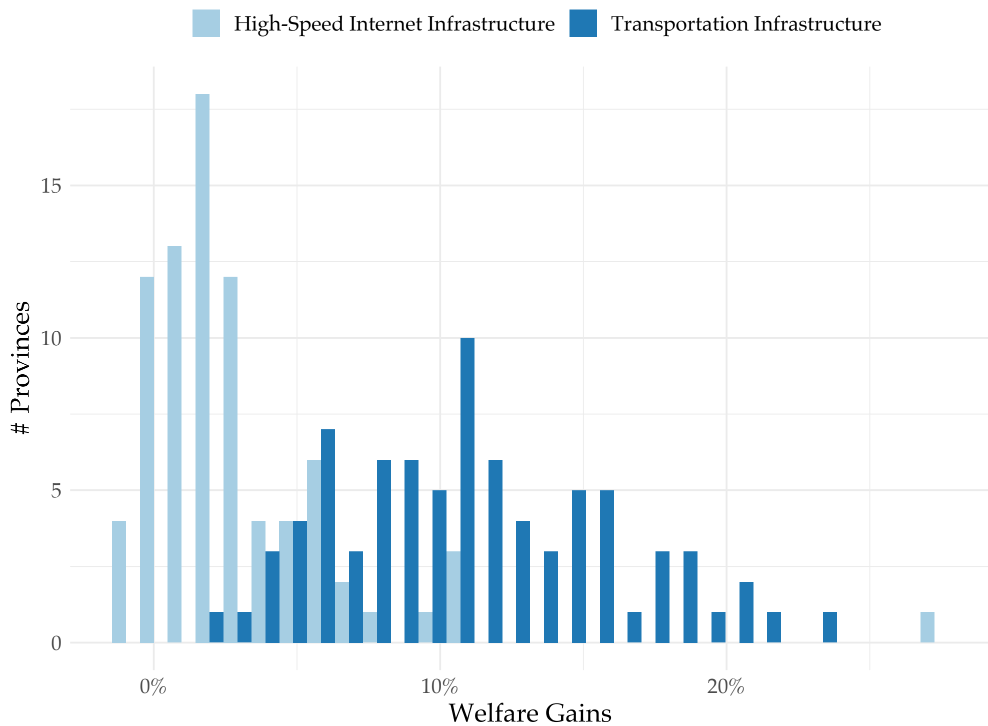
Turkish provinces. We observe significant heterogeneity in the gains across the provinces. The summary statistics indicate that the median gain is 2%, with an interquartile range of 1% to 4%. However, the maximum gain is as high as 27% for Istanbul, the most populous province in Turkey.[5] We also observe that the provinces with a higher percent increase in fibre length gained more, as shown in Figure 7. It is worth noting that this is the first study to assess the welfare gains resulting from ICT infrastructure improvements in a general equilibrium setting.

As such, we compare our findings with those of transportation infrastructure improvements made between 2005 and 2010. Turkey implemented an extensive programme in the 2000s to expand the lane capacity of existing roads. The goal was to improve safety and the reliability of travel times over the national transportation grid through investments across the country. Specifically, a substantial share of existing two-lane single-carriageways with two-way traffic was upgraded to dual carriageways separated by a small earthen medium, with two-lane one-way traffic on each carriageway.[6] To compare the two types of infrastructure improvements, we conduct a hypothetical experiment to determine the impact on real income if the 2012 Turkish economy had experienced similar transportation improvements as those between 2005 and 2010. The summary statistics in-

[5] See Figure A7 for details.

[6] See Cosar, Demir, Ghose, and Young (2021) for a detailed discussion.

Figure 8: **Welfare Effects of High-Speed Internet vs Transportation Infrastructure**



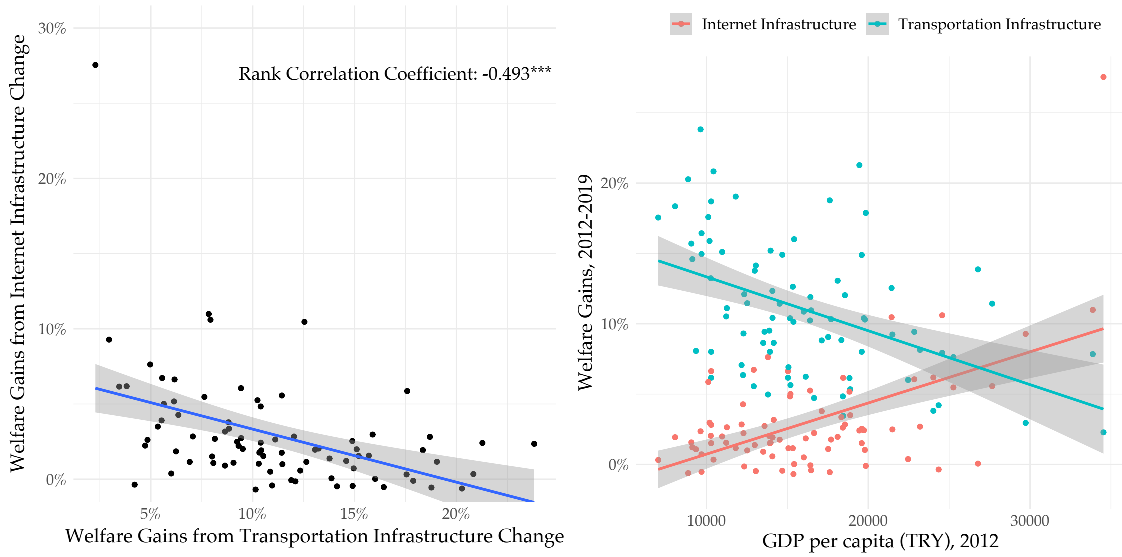
Note: This figure is a histogram of welfare changes across provinces. The lighter shade is for welfare changes due to internet infrastructure changes and the darker shade for transportation infrastructure changes.

dicate that the median gain resulting from transportation infrastructure improvements is 10%, with an interquartile range of 8% to 11% and a maximum gain of 24%. Interestingly, we observe that the gains from internet infrastructure improvements are negatively rank-correlated with the gains from transportation infrastructure improvements. Additionally, we note that while the poorer provinces gained more on average from transportation infrastructure improvements it is the richer provinces gained more on average from internet infrastructure improvements. This is illustrated in Figure 9. However, although the gains from internet infrastructure improvements are modest relative to those from transportation infrastructure, they are still of comparable magnitude.

8 Conclusion

This paper assesses the impact of high-speed internet infrastructure on economic growth through intra-national trade in input markets. Our analysis, which uses data on Turkish firm-to-firm linkages and the rollout of high-speed internet in Turkey, reveals that

Figure 9: **Welfare Effects of High-Speed Internet vs Transportation Infrastructure**



Note: This left panel plot welfare gains from internet infrastructure changes against that arising from transportation infrastructure changes across Turkish provinces. The right panel plot welfare changes from both internet and transportation infrastructure changes against per capita GDP of Turkish provinces.

the introduction of high-speed internet at both the origin and destination improves trade and encourages firms to diversify their input sourcing across a wider variety of suppliers. Furthermore, our model of spatial equilibrium featuring endogenous formation of input-output linkages between firms shows that high-speed internet infrastructure has heterogeneous but substantial welfare effects across the Turkish economy. The results of this study contribute to the growing body of literature on the importance of digital infrastructure in shaping economic outcomes and provide useful insights for policymakers looking to enhance economic growth and improve welfare through investments in high-speed internet infrastructure.

References

ACEMOGLU, D. AND P. D. AZAR (2020): "Endogenous Production Networks," *Econometrica*, 88, 33–82.

ACEMOGLU, D., V. M. CARVALHO, A. OZDAGLAR, AND A. TAHBAZ-SALEHI (2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80, 1977–2016.

AKERMAN, A., E. LEUVEN, AND M. MOGSTAD (2022): "Information Frictions, Internet, and the Relationship between Distance and Trade," *American Economic Journal: Applied Economics*, 14, 133–63.

AL-NAJJAR, N. I. (2004): "Aggregation and the Law of Large Numbers in Large Economies," *Games and Economic Behavior*, 47, 1 – 35.

ALLEN, T. (2014): "Information Frictions in Trade," *Econometrica*, 82, 2041–2083.

ALLEN, T. AND C. ARKOLAKIS (2014): " Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 129, 1085–1140.

ALLEN, T., C. ARKOLAKIS, AND X. LI (2020): "On the Equilibrium Properties of Network Models with Heterogeneous Agents," Working Paper 27837, National Bureau of Economic Research.

ALVAREZ, F. AND R. E. LUCAS (2007): "General Equilibrium Analysis of the Eaton-Kortum model of International Trade," *Journal of Monetary Economics*, 54, 1726 – 1768.

ARKOLAKIS, C., F. HUNEEUS, AND Y. MIYAUCHI (2023): "Spatial Production Networks," Working Paper 30954, National Bureau of Economic Research.

BAQAEE, D. R. AND E. FARHI (2019a): "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem," *Econometrica*, 87, 1155–1203.

——— (2019b): "Networks, Barriers, and Trade," Working Paper 26108, National Bureau of Economic Research.

——— (2020): "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 135, 105–163.

BERNARD, A. B., E. DHYNE, G. MAGERMAN, K. MANOVA, AND A. MOXNES (2022): "The Origins of Firm Heterogeneity: A Production Network Approach," *Journal of Political Economy*, 130, 1765–1804.

BERNARD, A. B., A. MOXNES, AND Y. U. SAITO (2019): "Production Networks, Geography, and Firm Performance," *Journal of Political Economy*, 127, 639–688.

——— (2020): "The Geography of Knowledge Production: Connecting Islands and Ideas," Working Paper, Dartmouth College.

BERTOLI, S., J. F.-H. MORAGA, AND L. GUICHARD (2020): "Rational Inattention and Migration Decisions," *Journal of International Economics*, 126, 103364.

BIGIO, S. AND J. LAO (2020): "Distortions in Production Networks," *The Quarterly Journal of Economics*, 135, 2187–2253.

BREGMAN, L. (1967): "The Relaxation Method of Finding the Common Point of Convex Sets and its Application to the Solution of Problems in Convex Programming," *USSR Computational Mathematics and Mathematical Physics*, 7, 200–217.

CALIENDO, L., F. PARRO, E. ROSSI-HANSBERG, AND P.-D. SARTE (2017a): "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," *The Review of Economic Studies*, 85, 2042–2096.

CHANEY, T. (2014): "The Network Structure of International Trade," *American Economic Review*, 104, 3600–3634.

COSAR, A. K., B. DEMIR, D. GHOSE, AND N. YOUNG (2021): "Road Capacity, Domestic Trade and Regional Outcomes," *Journal of Economic Geography*, 22, 901–929.

CRISTEA, A. D. (2011): "Buyer-Seller Relationships in International Trade: Evidence from U.S. States' Exports and Business-Class Travel," *Journal of International Economics*, 84, 207–220.

DASGUPTA, K. AND J. MONDRIA (2018): "Inattentive Importers," *Journal of International Economics*, 112, 150–165.

DEKLE, R., J. EATON, AND S. KORTUM (2008): "Global Rebalancing with Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers*, 55, 511–540.

DEMIR, B., A. C. FIELER, D. XU, AND K. K. YANG (2021): "O-Ring Production Networks," Working Paper 28433, National Bureau of Economic Research.

——— (2022): "Financial Constraints and Propagation of Shocks in Production Networks," *The Review of Economics and Statistics*, 1–46.

DICKSTEIN, M. J. AND E. MORALES (2018): "What do Exporters Know?" *The Quarterly Journal of Economics*, 133, 1753–1801.

DONALDSON, D. (2015): "The Gains from Market Integration," *Annual Review of Economics*, 7, 619–647.

——— (2018): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 108, 899–934.

EATON, J., S. S. KORTUM, AND F. KRAMARZ (2022): "Firm-to-Firm Trade: Imports, Exports, and the Labor Market," Working Paper 29685, National Bureau of Economic Research.

FERNANDES, A. M., A. MATTOO, H. NGUYEN, AND M. SCHIFFBAUER (2019): "The Internet and Chinese Exports in the Pre-Ali Baba Era," *Journal of Development Economics*, 138, 57–76.

FOSGERAU, M., E. MELO, A. DE PALMA, AND M. SHUM (2020): "Discrete Choice and Rational Inattention: A General Equivalence Result," *International Economic Review*, 61, 1569–1589.

FREUND, C. L. AND D. WEINHOLD (2004): "The Effect of the Internet on International Trade," *Journal of International Economics*, 62, 171–189.

HEAD, K. AND T. MAYER (2014): "Gravity Equations: Workhorse,Toolkit, and Cookbook," in *Handbook of International Economics*, ed. by G. Gopinath, E. Helpman, and K. Rogoff, Elsevier, vol. 4 of *Handbook of International Economics*, 131 – 195.

HJORT, J. AND J. POULSEN (2019): "The Arrival of Fast Internet and Employment in Africa," *American Economic Review*, 109, 1032–79.

HJORT, J. AND L. TIAN (2021): "The Economic Impact of Internet Connectivity in Developing Countries," Working Paper, INSEAD.

HOTZ, V. J. AND R. A. MILLER (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60, 497–529.

HUNEEUS, F. (2020): "Production Network Dynamics and the Propagation of Shocks," Working Paper, Central Bank of Chile.

JIANG, X. (2023): "Information and Communication Technology and Firm Geographic Expansion," Working Paper, University of California - Davis.

JUHÁSZ, R. AND C. STEINWENDER (2018): "Spinning the Web: The Impact of ICT on Trade in Intermediates and Technology Diffusion," NBER Working Papers 24590, National Bureau of Economic Research, Inc.

LIM, K. (2018): "Endogenous Production Networks and the Business Cycle," Working Paper, University of Toronto.

MALGOUYRES, C., T. MAYER, AND C. MAZET-SONILHAC (2021): "Technology-induced Trade Shocks? Evidence from Broadband Expansion in France," *Journal of International Economics*, 133, 103520.

MATEJKA, F. AND A. MCKAY (2015): "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105, 272–98.

MIYAUCHI, Y. (2023): "Matching and Agglomeration: Theory and Evidence from Japanese Firm-to-Firm Trade," Working Paper, Boston University.

OBERFIELD, E. (2018): "A Theory of Input-Output Architecture," *Econometrica*, 86, 559–589.

PANIGRAHI, P. (2022a): "Endogenous Spatial Production Networks: Quantitative Implications for Trade & Productivity," Working Paper, Johns Hopkins University.

——— (2022b): "Triad Trade & Small Worlds of Large Spatial Production Networks," Working Paper, Johns Hopkins University.

RAUCH, J. E. AND V. TRINDADE (2003): "Information, International Substitutability, and Globalization," *American Economic Review*, 93, 775–791.

REDDING, S. J. (2016): "Goods Trade, Factor Mobility and Welfare," *Journal of International Economics*, 101, 148 – 167.

——— (2022): "Trade and Geography," in *Handbook of International Economics: International Trade, Volume 5*, ed. by G. Gopinath, E. Helpman, and K. Rogoff, Elsevier, vol. 5 of *Handbook of International Economics*, 147–217.

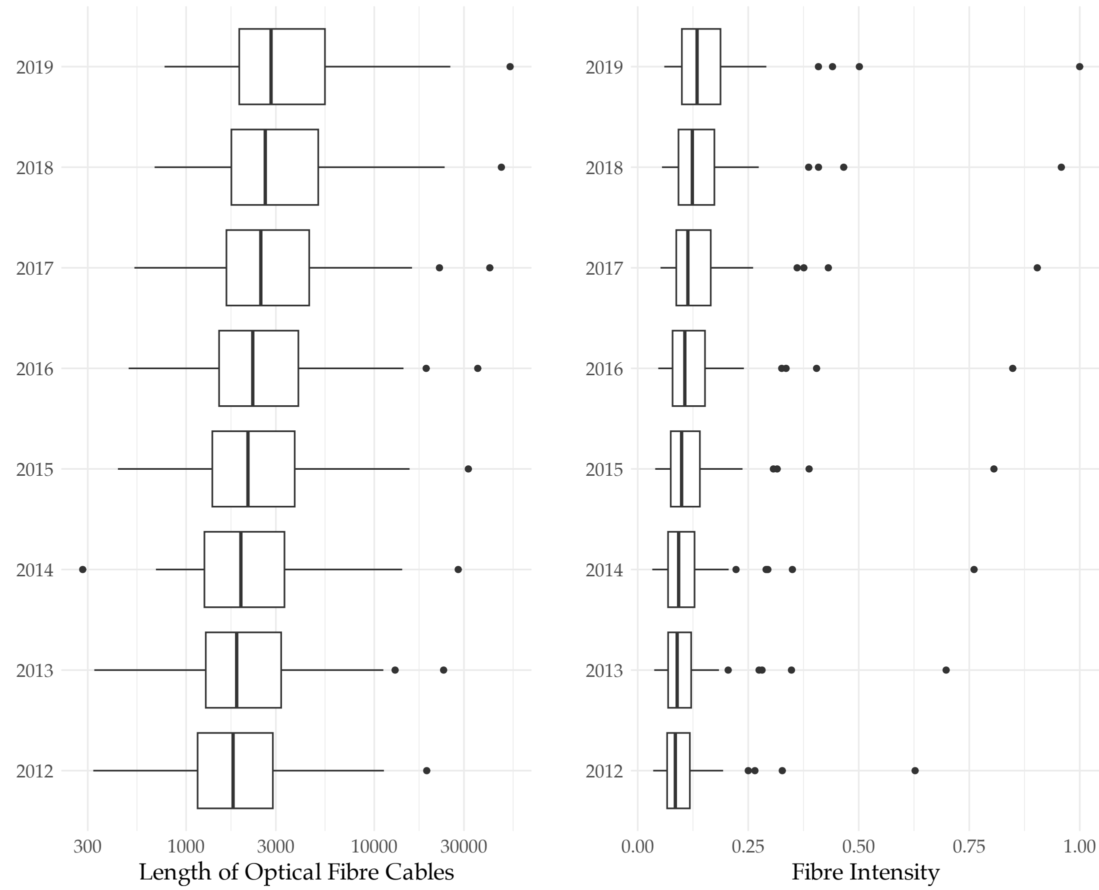
REDDING, S. J. AND E. ROSSI-HANSBERG (2017): "Quantitative Spatial Economics," *Annual Review of Economics*, 9, 21–58.

SHANNON, C. E. (1948): "A Mathematical Theory of Communication," *Bell System Technical Journal*, 27, 379–423.

SIMS, C. A. (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50, 665–690.

A Appendix: Background & Data

Figure A1: **Change in fibre cable length between 2012-2019**



Note: The left panel is a box and whiskers plot of length of optical fibre cables rolled out across Turkish provinces over years. The right panel is a box and whiskers plot of fibre intensity across Turkish provinces over the years.

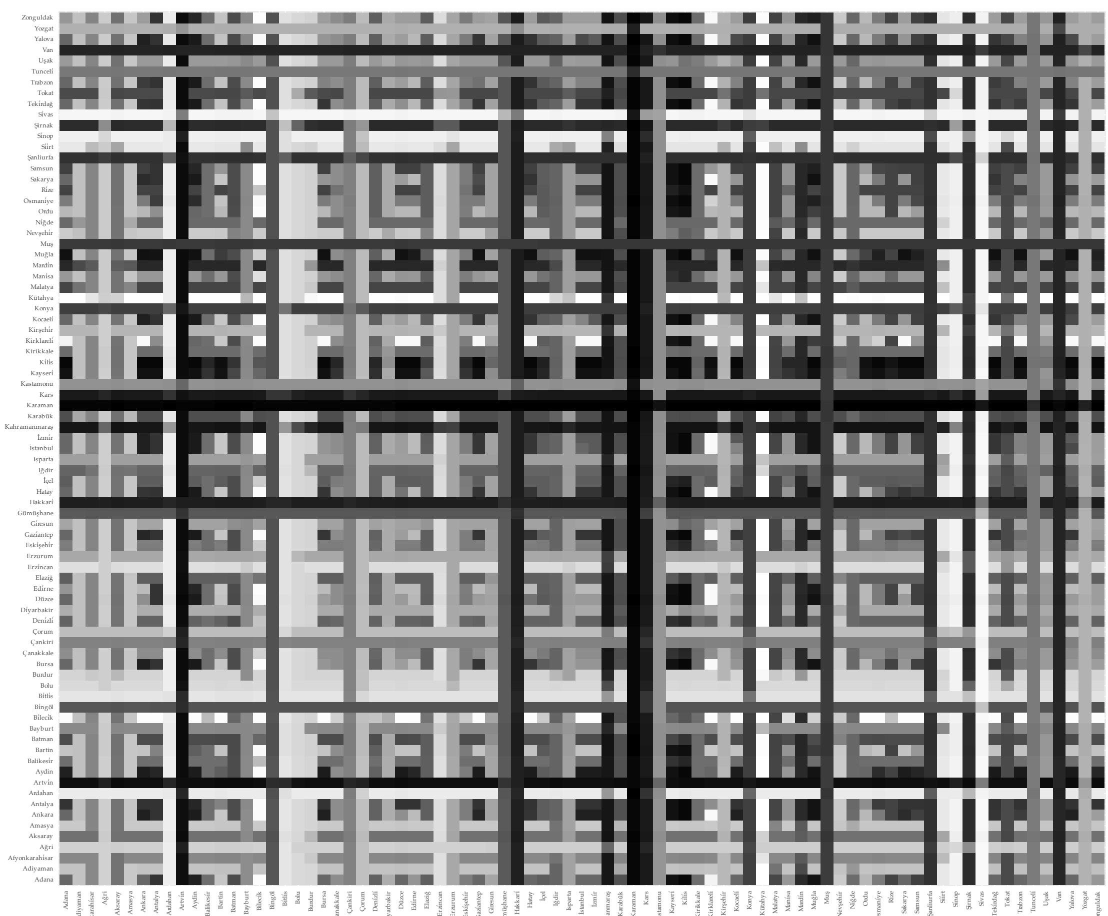
Table A1: **Firms' adoption of high-speed internet**

	(1)	(2)
Fibre Intensity	1.660**	
	(0.648)	
Fibre Intensity ×		
Food, beverages, and tobacco		1.648**
		(0.661)
Textiles, clothing, and footwear		1.637**
		(0.647)
Wood and paper products		1.666**
		(0.657)
Coke, petroleum, chemical products, and pharmaceuticals		1.668**
		(0.648)
Plastics and non-metallic mineral products		1.683**
		(0.644)
Basic metals		1.361**
		(0.651)
Fabricated metal products and general-purpose machinery		1.611**
		(0.644)
Computer, electronic, electrical and optical products		1.713**
		(0.665)
Manufacture of motor vehicles and ships		1.603**
		(0.651)
Furniture and other manufacturing		1.736**
		(0.650)
Trade		1.614**
		(0.643)
Professional services		1.629**
		(0.650)
Observations	3337	3337

Note: * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered by province, in parentheses. All columns include firm size (measured in terms of employment) as a control. The corresponding specification is in (1). Column (1) presents results at the province-year level and column (2) at the province-firm size-year level. The latter includes dummies for each firm size.

B Appendix: Empirical Evidence

Figure A2: **Change in Bilateral Fibre Connectivity**

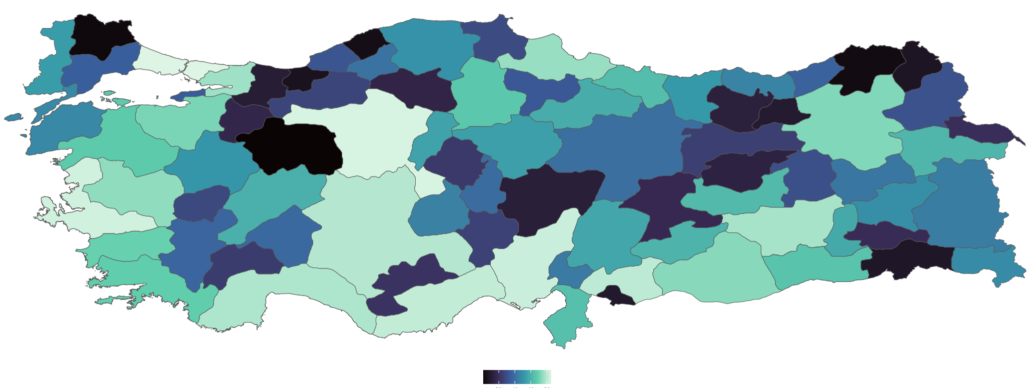


Note: This figure depicts the change in fibre connectivity between each pair of Turkish provinces during the period 2012-2019. Fibre Connectivity is measure as the minimum of fibre intensity between provinces in a pair. Darker shades represent higher values.

Figure A3: **BOTAS Oil and Gas Pipeline Network**

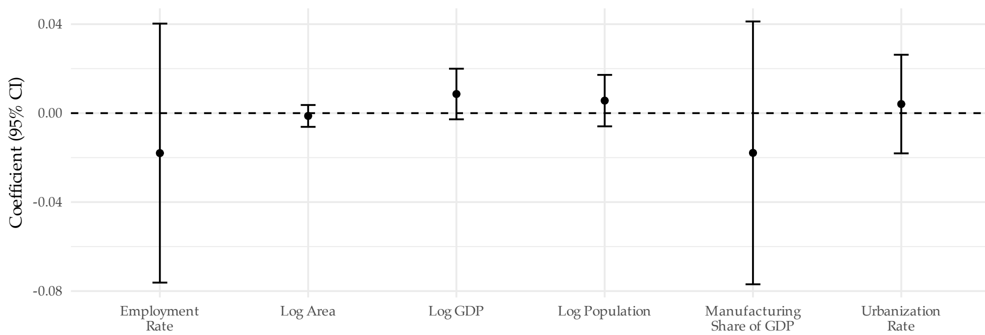
Note: This map shows the gas pipeline network of BOTAS.

Figure A4: **Distance to BOTAS Pipeline Network**



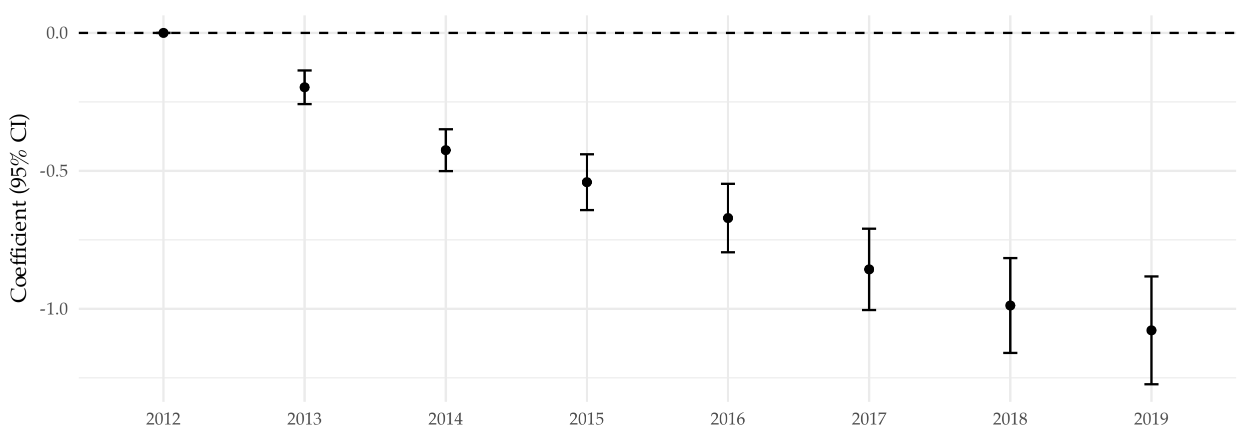
Note: This map depicts the spatial variation in weighted distance of provinces to the BOTAS pipeline network. The distance of a province to BOTAS pipeline is calculated as the weighted average of the shortest distance of its districts to the pipeline where each district is weighted by its population. Lighter shades represent higher percentiles.

Figure A5: **Distance to BOTAS pipelines and Initial Province Characteristics**



Note:See Section 3.3 for discussion. The distance of a province to BOTAS pipelines is constructed as the weighted average of the distances of districts within the province where the district population are used as weights. This figure plots the coefficient estimates and the corresponding 95% confidence intervals obtained from regressing this distance on initial provincial characteristics (pertaining to 2011), controlling for NUTS2 level fixed effects.

Figure A6: **Distance to BOTAS pipelines and First Stage Estimates**



Note: See Section 3.3 for discussion. The distance of a province to BOTAS pipelines is constructed as the weighted average of the distances of districts within the province where the district population are used as weights. This figure plots the coefficient estimates and the corresponding 95% confidence intervals obtained from first-stage regression of fibre connectivity on distance to BOTAS pipelines interacted with year dummies.

Table A2: **Reallocation and Diversification in Input Sourcing: Additional Controls**

Dependent Variable:	Cost Share		No. Suppliers		Cost Share HHI		New Connections	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fibre Connectivity	0.467***	0.515***	0.291***	0.323***	-0.0916***	-0.105***	0.0531***	0.0549***
	(0.0559)	(0.0590)	(0.0316)	(0.0280)	(0.0129)	(0.0144)	(0.0198)	(0.0196)
Difference in GDP p.c.	0.208**		0.165***		-0.0735***		0.0483**	
	(0.0808)		(0.0393)		(0.0195)		(0.0242)	
Mobile Connectivity		-0.0615		0.0255		-0.0174		0.0981
		(0.231)		(0.105)		(0.0626)		(0.0667)
Fixed Effects:								
Buyer×Year	✓	✓	✓	✓	✓	✓	✓	✓
Origin×Year	✓	✓	✓	✓	✓	✓	✓	✓
Origin×Destination	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,230,473	2,230,473	2,230,473	2,230,473	2,230,473	2,230,473	2,230,473	2,230,473

Note: See Section 3.2 for discussion. Each observation pertains to a buyer firm, an origin province and a year. All variables are in natural logarithms. No. Suppliers is the number of suppliers of the buyer firm located in a given origin province. Cost share is the fraction of purchases of a buyer from a given source province. Cost Share HHI is the Herfindahl-Hirschman Index of cost shares of suppliers of a buyer, which are located in a given origin province. New Connections is the number of new suppliers relative to the year before. For a pair of origin and destination provinces in a given year, difference in GDP p.c. is the absolute difference in GDP per capita and mobile connectivity is computed as the minimum of 3G/4G mobile subscribers per capita between both provinces. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

Table A3: **Reallocation and Diversification in Input Sourcing: Alternative Measure of Bilateral Connectivity**

Dependent Variable:	Cost Share (1)	No. Suppliers (2)	Cost Share HHI (3)	New Connections (4)
Fibre Connectivity, Alternative	0.255***	0.163***	-0.0534***	0.0316***
	(0.0274)	(0.0148)	(0.00682)	(0.00981)
Fixed Effects:				
Buyer×Year	✓	✓	✓	✓
Origin×Year	✓	✓	✓	✓
Origin×Destination	✓	✓	✓	✓
Observations	2,230,473	2,230,473	2,230,473	2,230,473

Note: See 3.2 for discussion. Each observation pertains to a buyer firm, an origin province and a year. All variables are in natural logarithms. No. Suppliers is the number of suppliers of the buyer firm located in a given origin province. Cost share is the fraction of purchases of a buyer from a given source province. Cost Share HHI is the Herfindahl-Hirschman Index of cost shares of suppliers of a buyer, which are located in a given origin province. New Connections is the number of new suppliers relative to the year before. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

Table A4: **Reallocation and Diversification in Input Sourcing: Placebo Test**

Dependent Variable:	Cost Share (1)	No. Suppliers (2)	Cost Share HHI (3)	New Connections (4)
Cable TV Connectivity	-2.210	-0.375	-0.0624	-0.315
	(1.754)	(0.906)	(0.441)	(0.487)
Fixed Effects:				
Buyer×Year	✓	✓	✓	✓
Origin×Year	✓	✓	✓	✓
Origin×Destination	✓	✓	✓	✓
Observations	2,230,473	2,230,473	2,230,473	2,230,473

Note: See Section 3.2 for discussion. Each observation pertains to a buyer firm, an origin province and a year. All variables are in natural logarithms. Cable TV connectivity is computed as the minimum of cable TV subscribers per capita between both provinces. No. Suppliers is the number of suppliers of the buyer firm located in a given origin province. Cost share is the fraction of purchases of a buyer from a given source province. Cost Share HHI is the Herfindahl-Hirschman Index of cost shares of suppliers of a buyer, which are located in a given origin province. New Connections is the number of new suppliers relative to the year before. All columns also include interactions of bilateral travel time between source and destination with annual dummy variables. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

Table A5: **Reallocation and Diversification in Input Sourcing: Including Non-Manufacturing Suppliers**

Dependent Variable:	Cost Share (1)	No. Suppliers (2)	Cost Share HHI (3)	New Connections (4)
Fibre Connectivity	0.691***	0.398***	-0.122***	0.145***
	(0.0613)	(0.0338)	(0.0129)	(0.0182)
Fixed Effects:				
Buyer×Year	✓	✓	✓	✓
Origin×Year	✓	✓	✓	✓
Origin×Destination	✓	✓	✓	✓
Observations	3,362,435	3,362,435	3,362,435	3,362,435

Note: See Section 3.2 for discussion. Each observation pertains to a buyer firm, an origin province and a year. All variables are in natural logarithms. No. Suppliers is the number of suppliers of the buyer firm located in a given origin province. Cost share is the fraction of purchases of a buyer from a given source province. Cost Share HHI is the Herfindahl-Hirschman Index of cost shares of suppliers of a buyer, which are located in a given origin province. New Connections is the number of new suppliers relative to the year before. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

Table A6: **Reallocation and Diversification in Input Sourcing: Source of Variation in Fibre Intensity**

Dependent Variable:	Cost Share (1)	No. Suppliers (2)	Cost Share HHI (3)	New Connections (4)
Panel A				
Origin Fibre Intensity ×				
Destination Fibre Intensity in:				
2^{nd} Quartile	-0.121	-0.155	0.168	-0.0501
	(0.477)	(0.218)	(0.118)	(0.139)
3^{rd} Quartile	0.213	0.172**	-0.0735*	0.0119**
	(0.173)	(0.0709)	(0.0387)	(0.0463)
4^{th} Quartile	0.519***	0.331***	-0.108***	0.0637***
	(0.0541)	(0.0295)	(0.0137)	(0.0196)
Panel B				
Destination Fibre Intensity ×				
Origin Fibre Intensity in:				
2^{nd} Quartile	-0.267	-0.238	0.0933	-0.139
	(0.443)	(0.123)	(0.209)	(0.143)
3^{rd} Quartile	0.181	0.173**	-0.0483	0.0191
	(0.159)	(0.0695)	(0.0395)	(0.0485)
4^{th} Quartile	0.517***	0.330***	-0.109***	0.0649***
	(0.0549)	(0.0296)	(0.0137)	(0.0192)
Fixed Effects:				
Buyer×Year	✓	✓	✓	✓
Origin×Year	✓	✓	✓	✓
Origin×Destination	✓	✓	✓	✓
Observations	2,230,473	2,230,473	2,230,473	2,230,473

Note: See Section 3.2 for discussion. Each observation pertains to a buyer firm, an origin province and a year. All variables are in natural logarithms. No. Suppliers is the number of suppliers of the buyer firm located in a given origin province. Cost share is the fraction of purchases of a buyer from a given source province. Cost Share HHI is the Herfindahl-Hirschman Index of cost shares of suppliers of a buyer, which are located in a given origin province. New Connections is the number of new suppliers relative to the year before. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

Table A7: **Reallocation and Diversification in Input Sourcing: Controlling for Travel Time Interactions**

Dependent Variable:	Cost Share (1)	No. Suppliers (2)	Cost Share HHI (3)	New Connections (4)
Fibre Connectivity	0.288***	0.183***	-0.0653***	0.0530***
	(0.0501)	(0.0255)	(0.0137)	(0.0199)
Fixed Effects:				
Buyer×Year	✓	✓	✓	✓
Origin×Year	✓	✓	✓	✓
Origin×Destination	✓	✓	✓	✓
Observations	2,230,473	2,230,473	2,230,473	2,230,473

Note: See Section 3.2 for discussion. Each observation pertains to a buyer firm, an origin province and a year. All variables are in natural logarithms. No. Suppliers is the number of suppliers of the buyer firm located in a given origin province. Cost share is the fraction of purchases of a buyer from a given source province. Cost Share HHI is the Herfindahl-Hirschman Index of cost shares of suppliers of a buyer, which are located in a given origin province. New Connections is the number of new suppliers relative to the year before. All columns also include interactions of bilateral travel time between origin and destination with annual dummy variables. Travel time is calculated using the road network in 2010. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

Table A8: **Reallocation and Diversification in Input Sourcing: Excluding Origin-Destination Fixed Effects**

Dependent Variable:	Cost Share (1)	No. Suppliers (2)	Cost Share HHI (3)	New Connections (4)
Fibre Connectivity	0.485***	0.264***	-0.114***	0.140
	(0.155)	(0.115)	(0.056)	(0.092)
Travel Time	-0.224***	-0.198***	0.102***	-0.107
	(0.019)	(0.015)	(0.008)	(0.011)
Fixed Effects:				
Buyer×Year	✓	✓	✓	✓
Origin×Year	✓	✓	✓	✓
Observations	3,362,435	3,362,435	3,362,435	3,362,435

Note: See Section 3.2 for discussion. Each observation pertains to a buyer firm, an origin province and a year. All variables are in natural logarithms. No. Suppliers is the number of suppliers of the buyer firm located in a given origin province. Cost share is the fraction of purchases of a buyer from a given source province. Cost Share HHI is the Herfindahl-Hirschman Index of cost shares of suppliers of a buyer, which are located in a given origin province. New Connections is the number of new suppliers relative to the year before. All columns also include interactions of bilateral travel time between origin and destination with annual dummy variables. Travel time is calculated using the road network in 2010. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

Table A9: **Reallocation and Diversification in Input Sourcing: Interaction between Travel Time and Connectivity**

Dependent Variable:	Cost Share (1)	No. Suppliers (2)	Cost Share HHI (3)	New Connections (4)
Fibre Connectivity	0.294***	0.176***	-0.056***	0.0319
	(0.056)	(0.025)	(0.013)	(0.0197)
Travel Time × Fibre Connectivity	-0.093***	-0.064***	0.022***	-0.0134**
	(0.018)	(0.0095)	(0.0039)	(0.0056)
Fixed Effects:				
Buyer×Year	✓	✓	✓	✓
Origin×Year	✓	✓	✓	✓
Origin×Destination	✓	✓	✓	✓
Observations	3,362,435	3,362,435	3,362,435	3,362,435

Note: Each observation pertains to a buyer firm, an origin province and a year. All variables are in natural logarithms. No. Suppliers is the number of suppliers of the buyer firm located in a given origin province. Cost share is the fraction of purchases of a buyer from a given source province. Cost Share HHI is the Herfindahl-Hirschman Index of cost shares of suppliers of a buyer, which are located in a given origin province. New Connections is the number of new suppliers relative to the year before. Travel time is calculated using the road network in 2010. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

C Appendix: Estimation & Results

C.1 Proof of Proposition 1

Since the exact realization of match-specific productivities still remain unknown at supplier choice, using properties of the Fréchet distribution, we can reformulate the firm's problem described in equation (3) as:

$$\max_{\pi_d^0(b,k)} \left\{ -\mathbb{E}_\nu \left[\zeta \ln c_o(s) + \zeta \ln \tau_{od} - \ln \phi_{od} \right] - \zeta \kappa_d(b,k) \right\}. \tag{21}$$

The result then follows from Proposition 4 in Fosgerau, Melo, de Palma, and Shum (2020).

C.2 Proof of Proposition 2

The log-likelihood function for the inner nest is given by

$$\begin{aligned}
\mathcal{L}_{\text{inner}} \propto & \left(\sum_{b \in \mathcal{M}} \sum_{s \in \mathcal{M}} \pi_{od}(s \mid o, b) \ln \left(\bar{\pi}_{od}\left(s \mid o, -\right) \tilde{c}_o(s)^{-\zeta/\lambda_{od}} \right) \right) \\
& - \left(\sum_{b \in \mathcal{M}} \sum_{s \in \mathcal{M}} \pi_{od}(s \mid o, b) \ln \left(\sum_{s' \in \mathcal{M}_o} \bar{\pi}_{od}\left(s' \mid o, -\right) \tilde{c}_o(s')^{-\zeta/\lambda_{od}} \right) \right) \\
= & \sum_{d \in \mathcal{J}} \left(\sum_{s \in \mathcal{M}} \left(\sum_{b \in \mathcal{M}_d} \pi_{od}(s \mid o, b) \right) \ln \left(\bar{\pi}_{od}\left(s \mid o, -\right) \tilde{c}_o(s)^{-\zeta/\lambda_{od}} \right) \right) \\
& - \sum_{d \in \mathcal{J}} \left(\sum_{s \in \mathcal{M}} \left(\sum_{b \in \mathcal{M}_d} \pi_{od}(s \mid o, b) \right) \ln \left(\sum_{s' \in \mathcal{M}_o} \bar{\pi}_{od}\left(s' \mid o, -\right) \tilde{c}_o(s')^{-\zeta/\lambda_{od}} \right) \right) \\
= & \sum_{d \in \mathcal{J}} \left(\sum_{s \in \mathcal{M}} \left(\sum_{b \in \mathcal{M}_d} \pi_{od}(s \mid o, b) \right) \ln \left(\bar{\pi}_{od}\left(s \mid o, -\right) \tilde{c}_o(s)^{-\zeta/\lambda_{od}} \right) \right) \\
& - \sum_{d \in \mathcal{J}} M_d \left(\sum_{o \in \mathcal{J}} \ln \left(\sum_{s' \in \mathcal{M}_o} \bar{\pi}_{od}\left(s' \mid o, -\right) \tilde{c}_o(s')^{-\zeta/\lambda_{od}} \right) \right) \\
= & \sum_{d \in \mathcal{J}} \left(\sum_{s \in \mathcal{M}} \pi_{od}(s \mid o, \bullet) \ln \left(\bar{\pi}_{od}\left(s \mid o, -\right) \tilde{c}_o(s)^{-\zeta/\lambda_{od}} \right) \right) \\
& - \sum_{d \in \mathcal{J}} M_d \left(\sum_{o \in \mathcal{J}} \ln \left(\sum_{s' \in \mathcal{M}_o} \bar{\pi}_{od}\left(s' \mid o, -\right) \tilde{c}_o(s)^{-\zeta/\lambda_{od}} \right) \right)
\end{aligned}$$

In the above expression, $\pi_{od}\left(s \mid o, b\right) = \frac{\pi_{od}(s,b)}{\sum_{s' \in \mathcal{M}_o} \pi_{od}(s',b)}$ and $\pi_{od}\left(s \mid o, \bullet\right) = \sum_{b \in \mathcal{M}_d} \pi_{od}\left(s \mid o, b\right)$.

The first order condition with respect to $\tilde{c}_o(s)^{-\zeta}$ can be simplified as follows:

$$
\begin{aligned}
&\frac{\partial \mathcal{L}}{\partial \tilde{c}_o(s)^{-\zeta}} = 0 \\
\implies &\sum_{d\in\mathcal{J}} \pi_{od}(s \mid o, \bullet) \left(\frac{1/\lambda_{od}}{\tilde{c}_o(s)^{-\zeta}}\right) = \sum_{d\in\mathcal{J}} M_d \left(\frac{1/\lambda_{od}\bar{\pi}_{od}\left(s \mid o, -\right)\left(\tilde{c}_o(s)^{-\zeta}\right)^{1/\lambda_{od}-1}}{\sum_{s'\in\mathcal{M}_o}\bar{\pi}_{od}\left(s' \mid o, -\right)\tilde{c}_o(s')^{-\zeta/\lambda_{od}}}\right) \\
\implies &\sum_{d\in\mathcal{J}} \pi_{od}(s \mid o, \bullet)\left(1/\lambda_{od}\right) = \sum_{d\in\mathcal{J}} M_d \left(\frac{\bar{\pi}_{od}\left(s \mid o, -\right)\tilde{c}_o(s)^{-\zeta/\lambda_{od}}}{\sum_{s'\in\mathcal{M}_o}\bar{\pi}_{od}\left(s' \mid o, -\right)\tilde{c}_o(s')^{-\zeta/\lambda_{od}}}\right)\left(1/\lambda_{od}\right)
\end{aligned}
$$

This equation is satisfied for the following solution:

$$
\left(\frac{\bar{\pi}_{od}\left(s \mid o, -\right)\tilde{c}_o(s)^{-\zeta/\lambda_{od}}}{\sum_{s'\in\mathcal{M}_o}\bar{\pi}_{od}\left(s' \mid o, -\right)\tilde{c}_o(s')^{-\zeta/\lambda_{od}}}\right)^* = \frac{\pi_{od}(s \mid o, \bullet)}{M_d} \tag{22}
$$

For this to be the maximum likelihood estimate this solution must also simultaneously satisfy the first order condition with respect to λ_{od}. The first order conditions with respect to θ can be simplified as follows:

$$
\begin{aligned}
&\frac{\partial \mathcal{L}}{\partial \lambda_{od}} = 0 \\
\implies &\sum_{d\in\mathcal{J}}\left(\sum_{s\in\mathcal{M}} \pi_{od}(s \mid o, \bullet)\ln\left(\tilde{c}_o(s)^{-\zeta}\right)\left(-1/\lambda_{od}^2\right)\right) \\
&= \sum_{d\in\mathcal{J}} M_d\left(\sum_{o\in\mathcal{J}}\frac{\partial \ln\left(\sum_{s\in\mathcal{M}_o}\bar{\pi}_{od}\left(s \mid o, -\right)\tilde{c}_o(s)^{-\zeta/\lambda_{od}}\right)}{\partial \lambda_{od}}\right) \\
&= \sum_{d\in\mathcal{J}} M_d\left(\sum_{o\in\mathcal{J}}\sum_{s\in\mathcal{M}_o}\frac{\bar{\pi}_{od}\left(s \mid o, -\right)\tilde{c}_o(s)^{-\zeta/\lambda_{od}}\ln\tilde{c}_o(s)^{-\zeta}}{\sum_{s'\in\mathcal{M}_o}\bar{\pi}_{od}\left(s' \mid o, -\right)c_o(s')^{-1/\lambda_{od}}}\left(-1/\lambda_{od}^2\right)\right) \\
&= \sum_{d\in\mathcal{J}} M_d\left(\sum_{s'\in\mathcal{M}}\frac{\bar{\pi}_{od}\left(s \mid o, -\right)\tilde{c}_o(s)^{-\zeta/\lambda_{od}}\ln\tilde{c}_o(s)^{-\zeta}}{\sum_{s'\in\mathcal{M}_o}\bar{\pi}_{od}\left(s' \mid o, -\right)c_o(s')^{-1/\lambda_{od}}}\left(-1/\lambda_{od}^2\right)\right)
\end{aligned}
$$

Clearly, the above solution in equation (22) satisfies this first order condition.

The log-likelihood of the outer nest is proportional to:

$$
\mathcal{L}_{\text{outer}} \propto \sum_{s\in\mathcal{M}}\left(\sum_{b\in\mathcal{M}}\pi_{od}(s, b)\right)\ln\left(c_o^{-\zeta}\tau_{od}^{-\zeta}\phi_{od}\bar{\pi}_{od}\left(\bullet, -\right)\right)
$$

$$
-\sum_{d\in\mathcal{J}} M_d \ln\left(\sum_{o'\in\mathcal{J}} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}\phi_{o'd}\bar{\pi}_{o'd}(\bullet,-)\right)
$$

The likelihood equations for $\tau_{od}^{-\zeta}\phi_{od}\bar{\pi}_{od}(\bullet,-)$ are given by:

$$
\begin{aligned}
\frac{\left(\sum_{b\in\mathcal{M}_d}\sum_{s\in\mathcal{M}_o}\pi_{od}(s,b)\right)}{\tau_{od}^{-\zeta}\phi_{od}\bar{\pi}_{od}(\bullet,-)} &= \frac{M_d}{\sum_{o'} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}\phi_{o'd}\bar{\pi}_{o'd}(\bullet,-)} c_o^{-\zeta} \\
&= \frac{M_d}{\sum_{o'} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}\phi_{o'd}\bar{\pi}_{o'd}(\bullet,-)} c_o^{-\zeta} \\
\implies \tau_{od}^{-\zeta}\phi_{od}\bar{\pi}_{od}(\bullet,-) &= \frac{\left(\sum_{b\in\mathcal{M}_d}\sum_{s\in\mathcal{M}_o}\pi_{od}(s,b)\right)}{\frac{M_d}{\sum_{o'} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}\phi_{o'd}\bar{\pi}_{o'd}(\bullet,-)} c_o^{-\zeta}} \\
&= \frac{\left(\sum_{b\in\mathcal{M}_d}\pi_{od}(\bullet,b)\right)}{\frac{M_d}{\sum_{o'} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}\phi_{o'd}\bar{\pi}_{o'd}(\bullet,-)} c_o^{-\zeta}} \\
\left(\frac{c_o^{-\zeta}\tau_{od}^{-\zeta}\phi_{od}\bar{\pi}_{od}(\bullet,-)}{\sum_{o'} c_{o'}^{-\zeta}\tau_{o'd}^{-\zeta}\phi_{o'd}\bar{\pi}_{o'd}(\bullet,-)}\right)^{*} &= \frac{\sum_{b\in\mathcal{M}_d}\pi_{od}(\bullet,b)}{M_d}
\end{aligned}
$$

This gives us the desired result.

Table A10: **Model-Consistent Gravity Regressions: Robustness**

Dependent Variable:	Average Cost Share				
	(1)	(2)	(3)	(4)	(5)
Fibre Connectivity	0.344**	0.548***	0.593***		
	(0.147)	(0.122)	(0.119)		
(Alternative) Fibre Connectivity				0.297***	
				(0.059)	
Cable TV Connectivity					0.150
					(0.124)
Fixed Effects:					
Origin × Year	✓	✓	✓	✓	✓
Destination × Year	✓	✓	✓	✓	✓
Origin × Destination	✓	✓	✓	✓	✓
Controls:					
Travel Time × Year	✓				
Diff. in GDP p.c.		✓			
Mobile Connectivity			✓		
R^2	0.994	0.994	0.994	0.994	0.984
Observations	45,504	45,504	45,504	45,504	45,504

Note: See Section 5.3 for discussion. For a pair of origin and destination provinces in a given year, (a) Alternative Fibre Connectivity is the negative absolute difference in fibre intensity; (b) Difference in GDP p.c. is the absolute difference in GDP per capita; (c) Mobile Connectivity is computed as the minimum of 3G/4G mobile subscribers per capita between both provinces; (d) Cable TV connectivity is computed as the minimum of cable TV subscribers per capita between both provinces; and (e) Travel time is calculated using the road network in 2010. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

D Appendix: Aggregation

D.1 Continuum Approximation for Large Network Economies

The following definition formalizes the notion of the limiting economy in the context of this paper.

Definition 3. Consider a sequence of finite economies $\{\mathcal{E}_t : t \in \mathbb{N}\}$ where $\mathcal{E}_t \equiv \{\mathcal{M}_t, \mathcal{L}_t, \mathcal{J}_t\}$ is such that the t^{th} economy has the form $\mathcal{M}_t = \{m_1, \cdots, m_{M_t}\} \subset [0,1]$, $\mathcal{L}_t = \{\ell_1, \cdots, \ell_{L_t}\} \subset [0,1]$ and $\mathcal{J}_t = \mathcal{J}$. The uniform distribution on $\mathcal{M}_t$ is given by $\mathcal{U}_t^M\left(\mathcal{M}_t^0\right) = \frac{M_t^0}{M_t}$ for all $\mathcal{M}_t^0 \subset \mathcal{M}_t$. Similarly, the uniform distribution on $\mathcal{L}_t$ is given by $\mathcal{U}_t^L\left(\mathcal{L}_t^0\right) = \frac{L_t^0}{L_t}$ for all $\mathcal{L}_t^0 \subset \mathcal{L}_t$. Then, $\{\mathcal{E}_t : t \in \mathbb{N}\}$ is a discretizing sequence of economies if it satisfies:

1. $\mathcal{M}_t \subset \mathcal{M}_{t+1}$ and $\mathcal{L}_t \subset \mathcal{L}_{t+1}$ for all t,
2. $\lim_{t\to\infty} \mathcal{U}_t^M\left(\mathcal{M}_t \cap [a_l, a_h]\right) = \mathcal{U}\left([a_l, a_h]\right)$,
3. $\lim_{t\to\infty} \mathcal{U}_t^L\left(\mathcal{L}_t \cap [a_l, a_h]\right) = \mathcal{U}\left([a_l, a_h]\right)$,

where $\mathcal{U}\left(\bullet\right)$ denotes the uniform distribution with support over $[0,1]$ and $[a_l, a_h] \subset [0,1]$.

D.2 Proof of Proposition 4

The probability with which any firm at d sources from firms at o for any of its tasks is given by

$$\begin{aligned}
\pi_{od}^0\left(\bullet, -\right) &= \left(\lim_{t\to\infty} \frac{M_o}{M}\right)\left(\lim_{t\to\infty} \frac{1}{M_o} \sum_{s\in\mathcal{M}_o} \pi_{od}^0(s,-)\right) \\
&= \left(\lim_{t\to\infty} \frac{M_o}{M}\right)\left(\lim_{t\to\infty} \frac{1}{M_o} \sum_{s\in\mathcal{M}_o} \frac{c_o(s)^{-\zeta}\phi_{od}(s)\tau_{od}^{-\zeta}}{A_d}\right) \\
&= \frac{\mu_o \mathbb{E}\left[\phi_{od}(\cdot)\right] \mathbb{E}\left[c_o(\cdot)^{-\zeta}\right] \tau_{od}^{-\zeta}}{A_d} \\
&= \frac{\mu_o \overline{z_o^{\zeta}} w_o^{-\zeta(1-\alpha_o)} \mathbb{E}\left[\Gamma\left(1 - \frac{\alpha_o}{K_o(\cdot)}\right)^{K_o(\cdot)}\right] A_o^{\alpha_o} \overline{\phi_{od}} \tau_{od}^{-\zeta}}{A_d}
\end{aligned}$$

From Proposition 4 in Fosgerau, Melo, de Palma, and Shum (2020), it follows that $A_d = \sum_o \mu_o \tau_{od}^{-\zeta} \mathbb{E}\left[\phi_{od}\right] \mathbb{E}\left[c_o(\cdot)^{-\zeta}\right]$

$$
\begin{aligned}
c_o(\cdot) &= w_o^{1-\alpha_o}\left(\prod_{k\in\mathcal{K}_o(\cdot)} p_o(\cdot,k)^{1/K_o(\cdot)}\right)^{\alpha_o} \\
\implies \mathbb{E}\left[c_o(\cdot)^{-\zeta}\right] &= \mathbb{E}\left[\left(\frac{w_o^{1-\alpha_o}\left(\prod_{k\in\mathcal{K}_o(\cdot)} p_o(\cdot,k)^{1/K_o(\cdot)}\right)^{\alpha_o}}{z_o(\cdot)}\right)^{-\zeta}\right] \\
&= w_o^{-\zeta(1-\alpha_o)}\mathbb{E}\left[\prod_{k\in\mathcal{K}_o(\cdot)} p_o(\cdot,k)^{-\alpha_o\zeta/K_o(\cdot)}\right]\mathbb{E}\left[z_o(\cdot)^{\zeta}\right] \\
&= w_o^{-\zeta(1-\alpha_o)}\left(\mathbb{E}\left[\mathbb{E}\left[\prod_{k\in\mathcal{K}_o(\cdot)} p_o(\cdot,k)^{-\alpha_o\zeta/K_o(\cdot)}\mid K_o\right]\right]\right)\mathbb{E}\left[z_o(\cdot)^{\zeta}\right] \\
&= w_o^{-\zeta(1-\alpha_o)}\left(\mathbb{E}\left[\prod_{k\in\mathcal{K}_o(\cdot)}\Gamma\left(1-\frac{\alpha_o}{K_o(\cdot)}\right)A_o^{\frac{\alpha_o}{K_o(\cdot)}}\right]\right)\overline{z_o^{\zeta}} \\
&= \mathbb{E}\left[\Gamma\left(1-\frac{\alpha_o}{K_o(\cdot)}\right)^{K_o(\cdot)}\right]\overline{z_o^{\zeta}}w_o^{-\zeta(1-\alpha_o)}A_o^{\alpha_o}
\end{aligned}
$$

This implies that $\{A_d\}_{d\in\mathcal{J}}$ solves the following fixed point problem:

$$
A_d = \sum_o \mu_o\phi_{od}\tau_{od}^{-\zeta}\overline{z_o^{\zeta}}\mathbb{E}\left[\Gamma\left(1-\frac{\alpha_o}{K_o\left(\cdot\right)}\right)^{K_o(\cdot)}\right]w_o^{-\zeta(1-\alpha_o)}A_o^{\alpha_o}
$$

It can be similarly shown that effective prices for needs faced by households is also given by $F_{p_d}\left(\cdot\right)$ The following lemma states that the above fixed point problem that solves for market access is well-defined in the sense that it admits a unique positive solution. The proof strategy follows from Allen et al. (2020).

Lemma. *The following system of equations*

$$
\begin{aligned}
A_d &= \sum_o R_{od}A_o^{\alpha_o}, \\
R_{od} &= \mu_o\phi_{od}\tau_{od}^{-\zeta}\overline{z_o^{\zeta}}\mathbb{E}\left[\Gamma\left(1-\frac{\alpha_o}{K_o\left(\cdot\right)}\right)^{K_o(\cdot)}\right]w_o^{-\zeta(1-\alpha_o)}A_o^{\alpha_o}.
\end{aligned}
$$

1. *has at least one positive solution*

2. *has at most one positive solution (up to scale)*

3. *the unique solution can be computed as the limit of a simple iterative procedure.*

Proof. First, we establish existence of positive solution to the system of equations. Define operator $T : \mathbb{R}^J_{++} \to \mathbb{R}^J_{++}$ where $T(\boldsymbol{A}) = \left(\sum_o R_{o1} A_o^{\alpha_o}, \cdots, \sum_o R_{oJ} A_o^{\alpha_o}\right)'$. Note that all components of R_{od} are positive and finite. Then, by construction, for any d, not all R_{od}s are zero. Therefore, for any $\boldsymbol{A} \gg 0$, $\sum_o R_{o1} A_o^{\alpha_o} \geq \underline{A} > 0$. Further, there exists $\bar{A} < \infty$ such that $\sum_o R_{od} A_o^{\alpha_o} \leq \bar{A}$. Now consider the operator $T : \mathcal{A} \to \mathcal{A}$ defined by $T(A_1, \cdots, A_J) = \left(\sum_o R_{o1} A_o^{\alpha_o}, \cdots, \sum_o R_{oJ} A_o^{\alpha_o}\right)'$. Suppose $\mathcal{A} = \left\{\boldsymbol{A} \in \mathbb{R}^J_{++} \mid \underline{A} \leq A_d \leq \bar{A} \forall d\right\}$.Then, if $\boldsymbol{A} \gg 0$, it follows that $T(\boldsymbol{A}) \gg 0$. Note that $\mathcal{A}$ is closed and bounded. Since $\mathcal{A} \subset \mathbb{R}^J_{++}$, this implies that $\mathcal{A}$ is compact. Further, $\mathcal{A}$ is non-empty and convex, and T is continuous. Then, by Brouwer's fixed point theorem, $T(\bullet)$ has a fixed point. This establishes existence of a solution the system of equations.

To establish uniqueness, let's suppose by way of contradiction that the system of equations has two different solutions $\boldsymbol{A}^{(0)}, \boldsymbol{A}^{(1)}$ that are not linear transformations of each other. Denote $\bar{a} = \max_d \frac{A_d^{(1)}}{A_d^{(0)}}$ and $\underline{a} = \min_d \frac{A_d^{(1)}}{A_d^{(0)}}$. Notice that $\frac{\bar{a}}{\underline{a}} \geq 1$. Thus the system of equations can be expressed as:

$$\frac{A_d^{(1)}}{A_d^{(0)}} = \frac{\sum_o R_{od} \left(\frac{A_d^{(1)}}{A_d^{(0)}}\right)^{1-\alpha_o} \left(A_d^{(0)}\right)^{1-\alpha_o}}{A_d^{(0)}}$$

Suppose $\bar{d} = \arg\max_d \left(\frac{A_d^{(1)}}{A_d^{(0)}}\right)$ and $\underline{\alpha} = \min \alpha_o$, then we have:

$$\begin{aligned}
\frac{A_{\bar{d}}^{(1)}}{A_{\bar{d}}^{(0)}} &= \bar{a} \\
\implies \frac{\sum_o R_{o\bar{d}} \left(\frac{A_o^{(1)}}{A_o^{(0)}}\right)^{1-\alpha_o} \left(A_o^{(0)}\right)^{1-\alpha_o}}{A_{\bar{d}}^{(0)}} &= \bar{a} \\
\implies \frac{\sum_o R_{o\bar{d}} \bar{a}^{1-\underline{\alpha}} \left(A_o^{(0)}\right)^{1-\alpha_o}}{A_{\bar{d}}^{(0)}} &\geq M \\
\implies \frac{\sum_o R_{o\bar{d}} \left(A_o^{(0)}\right)^{1-\alpha_o}}{A_{\bar{d}}^{(0)}} \bar{a}^{1-\underline{\alpha}} &\geq \bar{a} \\
\implies \bar{a}^{\underline{\alpha}} &\leq 1
\end{aligned}$$

$$\implies \bar{a} \leq 1$$

Similarly, we can show that $\underline{a} \geq 1$. This implies that $\frac{\bar{a}}{\underline{a}} \leq 1$. But by construction $\frac{\bar{a}}{\underline{a}} \geq 1$. Therefore, it must be the case that $\frac{\bar{a}}{\underline{a}} = 1$ or $\boldsymbol{A}^{(0)} = \boldsymbol{A}^{(1)}$. This establishes uniqueness.

Next, we show that the solution to the system of equations can be obtained via a simple iterative procedure. Starting from any strictly positive $\boldsymbol{A}^{(0)}$, we construct a sequence $\boldsymbol{A}^{(t)}$ successively in the following way,

$$A_d^{(t)} = \sum_o R_{od} \left(A_o^{(t-1)} \right)^{\alpha_o}$$

Denote $\bar{a}^{(t)} = \max_d \frac{A_d^{(t)}}{A_d^{(t-1)}}$ and $\underline{a}^{(t)} = \min_d \frac{A_d^{(t)}}{A_d^{(t-1)}}$. Notice that $\frac{\bar{a}^{(t)}}{\underline{a}^{(t)}} \geq 1$.

Suppose $\bar{d} = \arg\max_d \left(\frac{A_d^{(t)}}{A_d^{(t-1)}} \right)$ and $\underline{\alpha} = \min \alpha_o$, then we have:

$$\begin{aligned} \frac{A_{\bar{d}}^{(t)}}{A_{\bar{d}}^{(t-1)}} &= \bar{a}^{(t)} \\ \implies \frac{\sum_o R_{o\bar{d}} \left(\frac{A_o^{(t-1)}}{A_o^{(t-2)}} \right)^{1-\alpha_o} \left(A_o^{(t-2)} \right)^{1-\alpha_o}}{A_{\bar{d}}^{(t-1)}} &= \bar{a}^{(t)} \\ \implies \frac{\sum_o R_{o\bar{d}} \left(A_o^{(0)} \right)^{1-\alpha_o}}{A_{\bar{d}}^{(0)}} \left(\bar{a}^{(t-1)} \right)^{1-\underline{\alpha}} &\geq \bar{a}^{(t)} \\ \implies \frac{\bar{a}^{(t)}}{\left(\bar{a}^{(t-1)} \right)^{1-\underline{\alpha}}} &\leq 1 \end{aligned}$$

Similarly, we can show that $\frac{\underline{a}^{(t)}}{\left(\underline{a}^{(t-1)} \right)^{1-\bar{\alpha}}} \geq 1$. This implies the following

$$\begin{aligned} \frac{\bar{a}^{(t)}}{\left(\bar{a}^{(t-1)} \right)^{1-\underline{\alpha}}} &\leq \frac{\underline{a}^{(t)}}{\left(\underline{a}^{(t-1)} \right)^{1-\bar{\alpha}}} \\ \implies \frac{\bar{a}^{(t)}}{\underline{a}^{(t)}} &\leq \frac{\left(\bar{a}^{(t-1)} \right)^{1-\underline{\alpha}}}{\left(\underline{a}^{(t-1)} \right)^{1-\bar{\alpha}}} \\ &\leq \frac{\left(\bar{a}^{(t-1)} \right)^{1-\underline{\alpha}}}{\left(\underline{a}^{(t-1)} \right)^{1-\underline{\alpha}}} \end{aligned}$$

$$\implies \frac{\bar{a}^{(t)}}{\underline{a}^{(t)}} \leq \frac{\bar{a}^{(t-1)}}{\underline{a}^{(t-1)}}$$

Since $\frac{\bar{a}^{(t)}}{\underline{a}^{(t)}} \geq 1 \forall t$, this implies that $\lim_{t\to\infty} \frac{\bar{a}^{(t)}}{\underline{a}^{(t)}} = 1$. That is, the solution can be computed as the limit of a simple iterative procedure. $\square$

D.3 Proof of Proposition 5

For any realization of σ, labor demand by firm b at d can be expressed as:

$$l_d(b,\sigma) = \frac{1}{w_d(\sigma)} (1-\alpha_d)\, c_d(b,\sigma) y_d(b,\sigma)$$

Substituting the above expression in the labor market clearing for location d, we obtain:

$$\begin{aligned} L_d &= \sum_{b\in\mathcal{M}_d} l_d(b,\sigma) \\ &= \sum_{b\in\mathcal{M}_d} \frac{1}{w_d(\sigma)} (1-\alpha_d)\, c_d(b,\sigma) y_d(b,\sigma) \\ \implies \sum_{b\in\mathcal{M}_d} c_d(b,\sigma) y_d(b,\sigma) &= \frac{w_d(\sigma) L_d}{1-\alpha_d} \end{aligned}$$

Goods market clearing condition for firm s located at o can be simplified as:

$$\begin{aligned} y_o(s,\sigma) &= \sum_d \sum_{b\in\mathcal{M}_d} \sum_{k\in\mathcal{K}_d(b)} \frac{\tau_{od}(s,\sigma) m_{od}(s,b,k,\sigma)}{a_{od}(s,b,k,\sigma)} \\ &\quad + \sum_d \sum_{i\in\mathcal{L}_d} \sum_{n\in\mathcal{N}_d(i)} \frac{\tau_{od}(s,\sigma) q_{od}(s,i,n,\sigma)}{g_{od}(s,i,n,\sigma)} \\ \implies c_o(s,\sigma) y_o(s,\sigma) &= \sum_d \alpha_d \sum_{b\in\mathcal{M}_d} \left(\frac{1}{K_d(b)} \sum_{k\in\mathcal{K}_d(b)} \mathbf{1}\left\{s = s_d^*(b,k,\sigma)\right\} \right) c_d(b,\sigma) y_d(b,\sigma) \\ &\quad + \sum_d \sum_{i\in\mathcal{L}_d} \left(\frac{1}{K_d(i)} \sum_{k\in\mathcal{K}_d(i)} \mathbf{1}\left\{s = s_d^*(i,k,\sigma)\right\} \right) (w_d(\sigma) + \Pi_d(\sigma)) \end{aligned}$$

$$\implies \underbrace{\sum_{s\in\mathcal{M}_o} c_o(s,\sigma)y_o(s,\sigma)}_{\text{(1) Supply}} = \underbrace{\sum_d \alpha_d \sum_{b\in\mathcal{M}_d}\left(\frac{1}{K_d(b)}\sum_{k\in\mathcal{K}_d(b)}\mathbf{1}\left\{s_d^*(b,k,\sigma)\in\mathcal{M}_o\right\}\right)c_d(b,\sigma)y_d(b,\sigma)}_{\text{(2) Intermediate Input Demand}}$$

$$+\underbrace{\sum_d\sum_{i\in\mathcal{L}_d}\left(\frac{1}{K_d(i)}\sum_{k\in\mathcal{K}_d(i)}\mathbf{1}\left\{s_d^*(i,k,\sigma)\in\mathcal{M}_o\right\}\right)(w_d(\sigma)+\Pi_d(\sigma))}_{\text{(3) Final Consumption Demand}}$$

We can simplify term (1) by making use of the labor market clearing condition as:

$$\begin{aligned}\text{Supply} &= \sum_{s\in\mathcal{M}_o} c_o(s,\sigma)y_o(s,\sigma)\\ &= \frac{w_o(\sigma)L_o}{1-\alpha_o}\end{aligned}$$

We can simplify term (2) as follows:

$$\begin{aligned}&\text{Intermediate Input Demand}\\ &= \sum_d \alpha_d \sum_{b\in\mathcal{M}_d}\left(\frac{1}{K_d(b)}\sum_{k\in\mathcal{K}_d(b)}\mathbf{1}\left\{s_d^*(b,k,\sigma)\in\mathcal{M}_o\right\}\right)c_d(b,\sigma)y_d(b,\sigma)\\ &= \sum_d \alpha_d \frac{\overbrace{\frac{1}{M_d}\sum_{b\in\mathcal{M}_d}\left(\frac{1}{K_d(b)}\sum_{k\in\mathcal{K}_d(b)}\mathbf{1}\left\{s_d^*(b,k,\sigma)\in\mathcal{M}_o\right\}\right)c_d(b,\sigma)y_d(b,\sigma)}^{(A)}}{\underbrace{\frac{1}{M_d}\sum_{b\in\mathcal{M}_d}c_d(b,\sigma)y_d(b,\sigma)}_{(B)}}\\ &\times \underbrace{\sum_{b\in\mathcal{M}_d}c_d(b,\sigma)y_d(b,\sigma)}_{=\frac{w_d(\sigma)L_d}{1-\alpha_d}}\end{aligned}$$

Term (A) can be simplified as follows:

$$
\begin{aligned}
(A) &= \frac{1}{M_d}\sum_{b\in\mathcal{M}_d}\left(\frac{1}{K_d(b)}\sum_{k\in\mathcal{K}_d(b)}\mathbf{1}\left\{s_d^*(b,k,\sigma)\in\mathcal{M}_o\right\}\right)c_d(b,\sigma)y_d(b,\sigma)\\
&\xrightarrow{t\to\infty}\mathbb{E}\left[\left(\frac{1}{K_d(\cdot)}\sum_{k\in\mathcal{K}_d(\cdot)}\mathbf{1}\left\{s_d^*(\cdot,k,\sigma)\in\mathcal{M}_o\right\}\right)c_d(\cdot,\sigma)y_d(\cdot,\sigma)\right]\\
&= \mathbb{E}\left[\left(\frac{1}{K_d(\cdot)}\sum_{k\in\mathcal{K}_d(\cdot)}\mathbf{1}\left\{s_d^*(\cdot,k,\sigma)\in\mathcal{M}_o\right\}\right)\right]\mathbb{E}\left[c_d(\cdot,\sigma)y_d(\cdot,\sigma)\right]\\
&= \mathbb{E}\left[\mathbb{E}\left[\left(\frac{1}{K_d(\cdot)}\sum_{k\in\mathcal{K}_d(\cdot)}\mathbf{1}\left\{s_d^*(\cdot,k,\sigma)\in\mathcal{M}_o\right\}\right)\mid K_d\right]\right]\mathbb{E}\left[c_d(\cdot,\sigma)y_d(\cdot,\sigma)\right]\\
&= \mathbb{E}\left[\frac{1}{K_d(\cdot)}\sum_{k\in\mathcal{K}_d(\cdot)}\mathbb{E}\left[\mathbf{1}\left\{s_d^*(\cdot,k,\sigma)\in\mathcal{M}_o\right\}\mid K_d\right]\right]\mathbb{E}\left[c_d(\cdot,\sigma)y_d(\cdot,\sigma)\right]\\
&= \mathbb{E}\left[\frac{1}{K_d(\cdot)}\sum_{k\in\mathcal{K}_d(\cdot)}\mathbb{E}\left[\mathbf{1}\left\{s_d^*(\cdot,\cdot,\sigma)\in\mathcal{M}_o\right\}\right]\right]\mathbb{E}\left[c_d(\cdot,\sigma)y_d(\cdot,\sigma)\right]\\
&= \mathbb{E}\left[\mathbf{1}\left\{s_d^*(\cdot,\cdot,\sigma)\in\mathcal{M}_o\right\}\right]\mathbb{E}\left[c_d(\cdot,\sigma)y_d(\cdot,\sigma)\right]\\
&= \mathbb{E}\left[\mathbf{1}\left\{s_d^*(\cdot,\cdot,\sigma)\in\mathcal{M}_o\right\}\right]\mathbb{E}\left[c_d(\cdot,\sigma)y_d(\cdot,\sigma)\right]\\
&= \pi_{od}\left(\bullet,-,\sigma_0\right)\mathbb{E}\left[c_d(\cdot,\sigma)y_d(\cdot,\sigma)\right]
\end{aligned}
$$

Term (B) can be simplified as follows:

$$
\begin{aligned}
(B) &= \frac{1}{M_d}\sum_{b\in\mathcal{M}_d}c_d(b,\sigma)y_d(b,\sigma)\\
&\xrightarrow{t\to\infty}\mathbb{E}\left[c_d(\cdot,\sigma)y_d(\cdot,\sigma)\right]
\end{aligned}
$$

Substituting (A) and (B) back in the Intermediate Input Demand,we obtain:

$$
\text{Intermediate Input Demand} = \sum_d \alpha_d\pi_{od}\left(\bullet,-,\sigma_0\right)\frac{w_d(\sigma)L_d}{1-\alpha_d}
$$

We can simplify term (3) as follows:

Final Consumption Demand

$$
\begin{aligned}
&= \sum_{d} \sum_{i \in \mathcal{L}_d} \left(\frac{1}{K_d(i)} \sum_{k \in \mathcal{K}_d(i)} \mathbf{1} \left\{ s_d^*(i,k,\sigma) \in \mathcal{M}_o \right\} \right) w_d(\sigma) \\
&= \sum_{d} \left(\frac{1}{L_d} \sum_{i \in \mathcal{L}_d} \left(\frac{1}{K_d(i)} \sum_{k \in \mathcal{K}_d(i)} \mathbf{1} \left\{ s_d^*(i,k,\sigma) \in \mathcal{M}_o \right\} \right) \right) w_d(\sigma) L_d \\
&\xrightarrow{t \to \infty} \sum_{d} \mathbb{E} \left[\frac{1}{K_d(i)} \sum_{k \in \mathcal{K}_d(i)} \mathbf{1} \left\{ s_d^*(i,k,\sigma) \in \mathcal{M}_o \right\} \right] w_d(\sigma) L_d \\
&= \sum_{d} \mathbb{E} \left[\mathbb{E} \left[\frac{1}{K_d(i)} \sum_{k \in \mathcal{K}_d(i)} \mathbf{1} \left\{ s_d^*(i,k,\sigma) \in \mathcal{M}_o \right\} \mid N_d \right] \right] w_d(\sigma) L_d \\
&= \sum_{d} \mathbb{E} \left[\frac{1}{K_d(i)} \sum_{k \in \mathcal{K}_d(i)} \mathbb{E} \left[\mathbf{1} \left\{ s_d^*(i,k,\sigma) \in \mathcal{M}_o \right\} \mid N_d \right] \right] w_d(\sigma) L_d \\
&= \sum_{d} \mathbb{E} \left[\frac{1}{K_d(i)} \sum_{k \in \mathcal{K}_d(i)} \mathbb{E} \left[\mathbf{1} \left\{ s_d^*(i,k,\sigma) \in \mathcal{M}_o \right\} \right] \right] w_d(\sigma) L_d \\
&= \sum_{d} \mathbb{E} \left[\mathbf{1} \left\{ s_d^*(\cdot,\cdot,\sigma) \in \mathcal{M}_o \right\} \right] w_d(\sigma) L_d \\
&= \sum_{d} \mathbb{E} \left[\mathbf{1} \left\{ s_d^*(\cdot,\cdot,\sigma) \in \mathcal{M}_o \right\} \right] w_d(\sigma) L_d \\
&= \sum_{d} \pi_{od} \left(\bullet, -, \sigma_0 \right) w_d(\sigma) L_d
\end{aligned}
$$

Putting these together we can further simplify the goods market clearing condition to obtain the desired result as follows:

$$
\begin{aligned}
& \frac{w_o(\sigma) L_o}{1-\alpha_o} = \sum_{d} \pi_{od}(\bullet, -, \sigma_0) \left(\frac{\alpha_d}{1-\alpha_d} + 1 \right) w_d(\sigma) L_d \\
\implies & \frac{w_o(\sigma) L_o}{1-\alpha_o} = \sum_{d} \pi_{od}(\bullet, -, \sigma_0) \frac{w_d(\sigma) L_d}{1-\alpha_d}
\end{aligned}
$$

Since $\{w_d(\sigma)\}_d$ solves the above system of equations for a given realization of σ_0, irrespective of the realization of σ_1, we conclude that $w_d(\sigma) = w_d(\sigma_0)$. That is, $\{w_d : d \in \mathcal{J}\}$ solves the following system of equations for given realization of σ_0, irrespective to realization of σ_1.

$$
\frac{w_o L_o}{1-\alpha_o} = \sum_{d} \pi_{od}(\bullet, -) \frac{w_d L_d}{1-\alpha_d}
$$

E Appendix: Quantitative Analysis

E.1 Expected Utility & Welfare Changes

Households residing at location d are heterogeneous both in their numbers of needs and match-specific taste shocks of using different suppliers' goods to fulfill their needs. Welfare at any location is then calculated in expectation. That is, $V_d = \mathbb{E}\left[V_d\left(\cdot\right)\right]$. With Cobb-Douglas utilities across tasks, indirect utility of household i residing at d is given by:

$$V_d(i) = \frac{w_d}{\prod_{k \in \mathcal{K}_d(i)} p_d(i,k)^{1/K_d(i)}}$$

Expected indirect utility of households at location d can then be derived as:

$$\begin{aligned}
V_d &= \mathbb{E}\left[V_d\left(\cdot\right)\right] \\
&= \mathbb{E}\left[w_d \prod_{k \in \mathcal{K}_d(\cdot)} p_d(\cdot,k)^{-1/K_d(\cdot)}\right] \\
&= w_d \mathbb{E}\left[\mathbb{E}\left[\prod_{k \in \mathcal{K}_d(\cdot)} p_d(\cdot,k)^{-1/K_d(\cdot)} \mid K_d\right]\right] \\
&= w_d \mathbb{E}\left[\prod_{k \in \mathcal{K}_d(\cdot)} \mathbb{E}\left[p_d(\cdot,\cdot)^{-1/K_d(\cdot)} \mid K_d\right]\right] \\
&= w_d \mathbb{E}\left[\prod_{k \in \mathcal{K}_d(\cdot)} \Gamma\left(1 - \frac{1}{\zeta K_d\left(\cdot\right)}\right) A_d^{\frac{1}{\zeta K_d(\cdot)}}\right] \\
&= \mathbb{E}\left[\Gamma\left(1 - \frac{1}{\zeta K_d\left(\cdot\right)}\right)^{K_d(\cdot)}\right] w_d A_d^{\frac{1}{\zeta}}
\end{aligned}$$

Welfare changes, i.e., changes in expected indirect utility at location d in response to shocks can be calculated as:

$$\widehat{V}_d = \widehat{w}_d \widehat{A}_d^{1/\zeta},$$

where $\widehat{w}_d$ denotes the change in wage and $\widehat{A}_d$ denotes change in market access at d.

E.2 Computation of Counterfactual Outcomes

For any change in σ_0, $\widehat{\boldsymbol{\delta}} \equiv \left\{\widehat{\delta}_{od} : (o,d) \in \mathcal{J} \times \mathcal{J}\right\}$, one can solve for change in wages $\widehat{\boldsymbol{w}} \equiv \{\widehat{w}_d : d \in \mathcal{J}\}$ with the following tâtonnement algorithm for some positive constant μ and tolerance value tol:

1. Start with a guess for the vector of change in wages, $\widehat{\boldsymbol{w}}^{(0)}$ and

2. For the vector of wage changes, in the t^{th} iteration $\widehat{\boldsymbol{w}}^{(t)}$, compute change in market access and endogenous trade costs as the solution to the following system of equations:

$$\widehat{A}_d^{(t)} = \sum_o \pi_{od} \widehat{\tau}_{od}^{-\zeta} \widehat{\phi}_{od} \left(\widehat{w}_o^{(t)}\right)^{-\zeta(1-\alpha_o)} \left(\widehat{A}_o^{(t)}\right)^{\alpha_o}$$

3. Compute counterfactual sourcing probabilities as:

$$\left(\pi_{od}^{(t)}\right)' = \pi_{od}^{(t)} \frac{\widehat{\tau}_{od}^{-\zeta} \widehat{\phi}_{od} \left(\widehat{w}_o^{(t)}\right)^{-\zeta(1-\alpha_o)} \left(\widehat{A}_o^{(t)}\right)^{\alpha_o}}{\widehat{A}_d^{(t)}}$$

4. Compute excess demand for labor $\boldsymbol{Z}\left(\widehat{\boldsymbol{w}}^{(t)}\right) \equiv \left\{Z_o\left(\widehat{\boldsymbol{w}}^{(t)}\right) : o \in \mathcal{J}\right\}$ as:

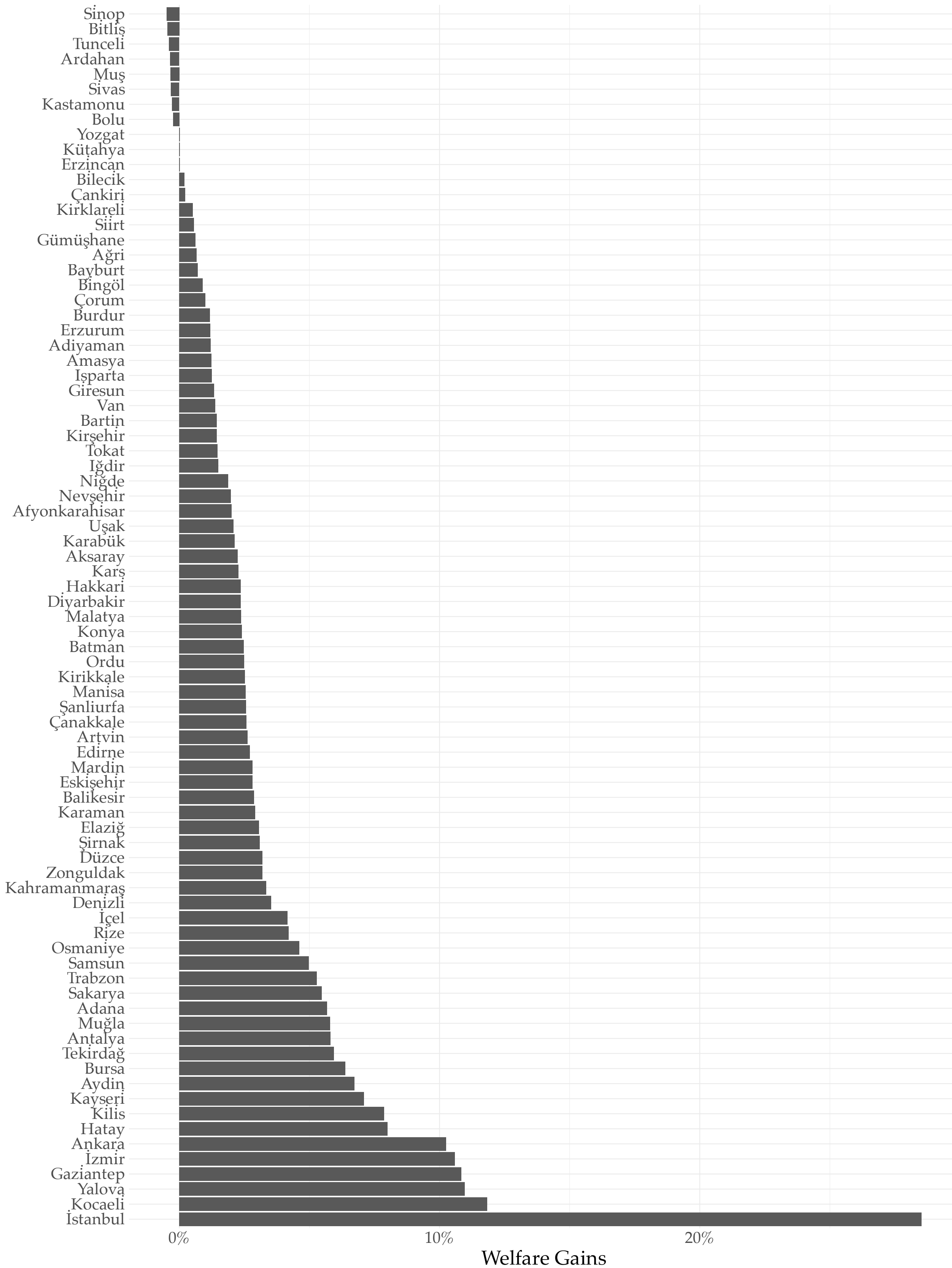
$$Z_o\left(\widehat{\boldsymbol{w}}^{(t)}\right) = \frac{1-\alpha_o}{w_o L_o} \sum_d \left(\pi_{od}^{(t)}\right)' \widehat{w}_d^{(t)} \frac{w_d L_d}{1-\alpha_d} - \widehat{w}_o$$

5. Update the vector of change in wages as $\widehat{\boldsymbol{w}}^{(t+1)} \leftarrow \widehat{\boldsymbol{w}}^{(t)} + \mu \boldsymbol{Z}\left(\widehat{\boldsymbol{w}}^{(t)}\right)$.

6. If $\|\widehat{\boldsymbol{w}}^{(t+1)} - \widehat{\boldsymbol{w}}^{(t)}\| > tol$, go back to (2), else end.

Welfare changes can then be computed as $\widehat{V}_d = \widehat{w}_d^{(\infty)} \left(\widehat{A}_d^{(\infty)}\right)^{\frac{1}{\zeta}}$.

Figure A7: **Welfare Effects of High-Speed Internet Infrastructure**



Note: This figure depicts the welfare impact of high-speed internet roll-out across Turkish provinces. The median Turkish province saw a 2% increase in welfare, with an interquartile range of 1% to 4%.