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

Banu Demir* Beata Javorcik† Piyush Panigrahi‡

March 2026

**Abstract**

Why is aggregate productivity lower in developing countries? This paper argues that a key part of the answer lies in information frictions that distort domestic production networks and hinder efficient supplier matching. It uses theory and empirics to study how such frictions constrain firm-to-firm trade and distinguishes two channels through which they operate: *communication frictions*, the cost of transmitting known information, and *information acquisition frictions*, the cost of discovering and evaluating suppliers. Combining firm-to-firm transaction data from Türkiye with rollout of fiber-optic cable, instrumented by proximity to a pre-existing infrastructure, we show that improved connectivity reshapes production networks along two margins: firms reallocate input purchases toward better-connected provinces and diversify their supplier base within those provinces. We develop and estimate a spatial model with endogenous production networks in which communication cost reductions drive reallocation across origins and information cost reductions broaden firms' supplier sets. Counterfactuals show that easing these frictions through fiber expansion generates sizeable welfare gains, highlighting the quantitative importance of information frictions for aggregate productivity.

JEL Codes: O33, L14, D85, D83, R12, F14, O18, L23.
Keywords: High-speed internet; fiber-optic infrastructure; supplier diversification; production networks; firm-to-firm transactions; rational inattention; digital infrastructure; network formation.
We thank our discussants Andrew Bernard, Julien Martin, Andreas Moxnes, and Lin Tian for their insightful comments. We also thank the audience at numerous seminars and conferences for their comments. We are grateful to the EIS team at the Ministry of Industry and Technology of the Republic of Türkiye for their support with the disclosure process, and Ookla for sharing with us their speed data for Türkiye. All interpretations, errors and omissions are the authors' own.

*Oxford, Bilkent, and CEPR; email: banu.demirpakel@economics.ox.ac.uk
†Oxford, EBRD, and CEPR; email: beata.javorcik@economics.ox.ac.uk
‡International Finance Corporation; email: ppanigrahi@ifc.org

## 1 Introduction

Why is aggregate productivity so much lower in developing countries than in advanced economies? A central explanation is high internal trade frictions that raise the cost of buying and selling goods across space and prevent an efficient allocation of resources within the domestic economy (Ramondo et al., 2016). Recent work has examined which frictions most constrain growth, highlighting the role of information frictions, particularly in limiting firms' access to input markets in developing countries (Atkin and Khandelwal, 2020). Firms in developing countries face elevated costs of identifying trading partners, verifying their quality, and coordinating transactions across space. Direct estimates of the welfare gains from alleviating these frictions, however, remain scarce. This paper fills this gap by distinguishing between two channels through which information frictions operate: *communication frictions*, the cost of transmitting known information between agents, and *information acquisition frictions*, the cost of discovering and processing information about potential trading partners.

It does so by focusing on the expansion of high-speed internet infrastructure, which offers a natural way to study how easing information frictions affects production networks. As firms increasingly rely on high-speed internet for real-time coordination with suppliers — through video conferencing, cloud-based data sharing, and digital file transfer — the availability and quality of digital infrastructure shape the spatial organization of sourcing relationships. Access to high-speed internet can influence firms' decisions through the same two channels identified above: it lowers the cost of coordinating with existing suppliers by improving communication quality and reliability, and it reduces the cost of discovering and evaluating new trading partners by facilitating information acquisition.

The paper makes several contributions to the existing literature. It provides the first general equilibrium quantification of welfare gains from ICT infrastructure that operate through firm-to-firm trade linkages, decomposing those gains into the distinct contributions of *communication costs* and *information acquisition costs*. This decomposition is not merely an accounting exercise: the two channels have different implications for which firms and regions benefit, how supply chains reorganize, and what kinds of complementary policies might amplify the returns to connectivity investment. Delivering on this requires two building blocks that are independently novel. On the empirical side, the paper provides causal evidence, using a novel instrument based on pre-existing pipeline infrastructure, that high-speed internet affects both the geography and concentration of input sourcing. On the theoretical side, it develops a tractable framework that embeds rational inattention into a model of endogenous production networks, allowing for separate identification of how internet connectivity lowers communication costs versus information acquisition costs.

We combine rich administrative microdata on firm-to-firm transactions in Türkiye with detailed information on the rollout of fiber-optic infrastructure across Turkish provinces between 2012 and 2019. Our empirical strategy exploits regulatory changes that generated plausibly exogenous variation in the spatial pattern of fiber deployment. In 2011, Türkiye deregulated its broadband market and permitted private internet service providers to utilize optical fiber cables previously laid alongside the national gas pipeline network. Because this pipeline infrastructure was designed decades earlier for energy transport rather than internet provision, and because proximity to the network determined the cost of subsequent fiber expansion, historical pipeline geography provides a powerful instrument for province-level internet connectivity.

Our main empirical findings establish two robust patterns. First, we find support for the *communication cost channel* being at work: firms reallocate input purchases toward provinces with better internet connectivity. A 10 percent increase in bilateral fiber connectivity raises the share of inputs sourced from a province by 4.3 percent. Second, conditional on sourcing from better-connected provinces, firms diversify more extensively across suppliers: they transact with more suppliers and distribute purchases more evenly and form new trading relationships at higher rates. The estimated magnitude of this *information acquisition cost channel* is economically meaningful: moving from the 10th to the 90th percentile of predicted connectivity reduces a firm's Herfindahl–Hirschman Index of purchase concentration by approximately 11 percent. All the patterns mentioned are robust to alternative measures of connectivity, controls for concurrent infrastructure investments and regional development, placebo tests, and event-study specifications that confirm the absence of pre-trends. We also find that diversification effects are stronger in industries with greater uncertainty about supplier quality, while no such effect is present for reallocation.

To rationalize these empirical patterns and quantify their aggregate implications, we develop a general equilibrium model of inter-provincial trade featuring endogenous supplier choice under information frictions. The model builds on Oberfield (2018) and Panigrahi (2021), incorporating rationally inattentive sourcing in the spirit of Dasgupta and Mondria (2018). Better internet connectivity operates through two channels. First, it raises the expected quality of match-specific productivity draws between buyers and suppliers, capturing the idea that reliable, high-bandwidth connections facilitate smoother coordination and reduce the friction of transacting across distance. This *communication channel* increases the probability of sourcing from well-connected provinces, generating the reallocation pattern observed in the data. Second, improved connectivity lowers the *cost of acquiring information* about potential suppliers. In our rational inattention framework, firms face a trade-off between precision (learning a lot about a few suppliers) and scope (learning a little about many). When information is expensive, firms concentrate attention on a small number of known partners; when information is cheap, they optimally spread attention more broadly, evaluating a larger pool of suppliers and distributing purchases more evenly across them. This information channel generates the diversification pattern that we document empirically.

The model yields supplier choice probabilities with a nested logit structure that maps directly into observed firm-to-firm trade data, providing a transparent link between theory and empirics. We estimate the key elasticities governing how connectivity affects each friction channel using a control function approach that leverages our instrumental variable. The calibrated model closely replicates observed changes in cost shares and supplier concentration between 2012 and 2019. In our model, information acquisition costs endogenously alter the elasticity of substitution across suppliers, breaking the constant elasticity structure that underpins standard sufficiency results for the class of trade models studied in Arkolakis et al. (2012).

Finally, we quantify general equilibrium welfare effects. The median Turkish province experienced a 0.36 percent annual increase in real income between 2012 and 2019 attributable to expansion of fiber infrastructure. Counterfactual decompositions indicate that reductions in *information acquisition costs* account for approximately 0.21 percentage points, while reductions in *communication costs* contribute about 0.13 percentage points. Welfare gains exhibit meaningful spatial heterogeneity, with initially remote or informationally disadvantaged provinces experiencing the largest benefits from improved connectivity.

In sum, our welfare calculations suggest that digital infrastructure investments generate substantial returns comparable to those stemming from physical infrastructure and operate through distinct mechanisms. These findings have immediate policy relevance. Governments worldwide are investing heavily in broadband infrastructure, with development banks prioritizing digital connectivity as a tool for economic growth. Our results suggest that such investments yield benefits not only through direct productivity gains but also through the reorganization of supply chains enabling firms to access better inputs and hedge against supply disruptions through diversification.

This paper relates to a large literature studying how internal trade frictions impede economic development. A central finding of this literature is that reductions in domestic trade costs whether through transportation infrastructure or market integration generate substantial welfare and productivity gains. Donaldson and Hornbeck (2016) and Donaldson (2018) estimate large gains from railroad expansion in the United States and India respectively. Allen and Arkolakis (2014) and Coşar et al. (2022) estimate gains from highway improvements in the United States and Türkiye respectively. Ramondo et al. (2016) show that domestic trade frictions can be quantitatively important in explaining cross-country income differences. Sotelo (2020) and Tombe and Zhu (2019) demonstrate that internal frictions generate spatial misallocation in agriculture and manufacturing. While much of this literature has focused on transportation costs as the primary friction, our paper shows that information frictions, mediated by digital infrastructure, operate through a distinct channel: they shape not only the volume of trade but the structure of production networks themselves, with aggregate implications that standard market-integration frameworks do not capture.

A second related literature studies the impact of information frictions on economic activity. On the empirical side, Jensen (2007) and Aker (2010) show that mobile phone adoption reduced price dispersion across spatially separated markets, Hjort and Poulsen (2019) study the employment effects of broadband internet in Africa, and Juhász and Steinwender (2018) demonstrates that the telegraph reduced the cost of transmitting price information with real effects on trade and production.[1] In a trade context, Rauch and Trindade (2003) show that ethnic networks facilitate trade by

[1] Other related work has also studied internet's effects on international trade flows (Freund and Weinhold, 2004; Fernandes et al., 2019; Malgouyres et al., 2021), labor markets (Akerman et al., 2015), firm-bank matching (Mazet-Sonilhac, 2021), and firm organization (Jiang, 2023). See Hjort and Tian (2025) for a comprehensive review.

reducing information frictions, Allen (2014) formalizes search frictions and estimates their welfare costs, and Dickstein and Morales (2018) provide evidence that information frictions distort firms' export decisions. On the theoretical side, Dasgupta and Mondria (2018) introduce rational inattention into a trade model where firms face costly information processing about foreign markets. A limitation of much of the empirical literature is that natural experiments in communication infrastructure simultaneously reduce both information acquisition and transmission costs, making it difficult to isolate the operative mechanism. We extend the rational inattention framework of Dasgupta and Mondria (2018) to allow for more flexible patterns of substitution across suppliers and link information costs directly to observable digital infrastructure. This enables both structural estimation of the distinct channels through which information frictions operate and quantification of their welfare implications.

Third, this paper contributes to a literature on production networks and their macroeconomic implications. Acemoglu et al. (2012) show that sectoral shocks can generate aggregate fluctuations when the input-output network is sufficiently asymmetric, while Baqaee and Farhi (2024) extend this framework to incorporate nonlinearities and misallocation. A growing body of work endogenizes the network structure itself: Chaney (2014), Oberfield (2018), and Demir et al. (2024) model how firms form linkages through search and learning, Eaton et al. (2022) and Arkolakis et al. (2025) study network formation in a multi-location general equilibrium framework. Our framework extends Panigrahi (2021) by introducing match-specific productivities that depend on internet connectivity and by incorporating rationally inattentive behavior into firms' supplier choice decisions. This allows us to capture the distinctive empirical pattern that connectivity affects not only which provinces firms source from but also how they allocate purchases across suppliers within those provinces.

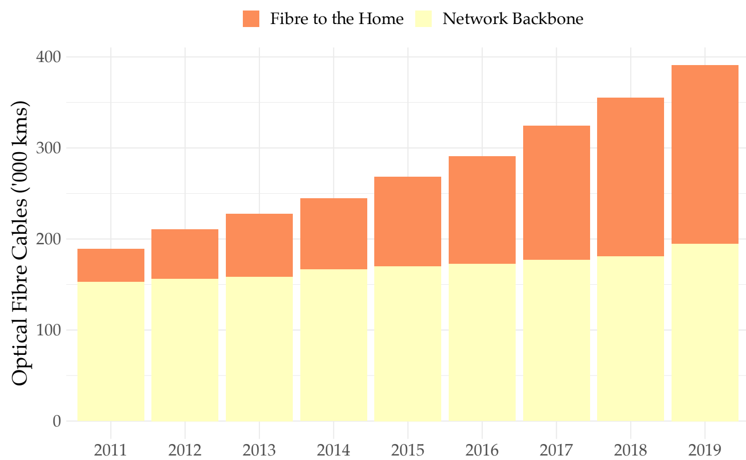
The remainder of the paper proceeds as follows. Section 2 describes the background of optical fiber cable rollout in Türkiye, the data sources, and the construction of our instrument. Section 3 presents the empirical evidence on the effects of fiber connectivity on firms' input sourcing patterns. Section 4 develops the model of input sourcing under rational inattention. Section 5 describes the estimation framework and presents structural parameter estimates. Section 6 reports the counterfactual welfare analysis. Section 7 concludes.

# 2 Background and Data

## 2.1 Rollout of Fiber Optic Infrastructure in Türkiye

Prior to 2010, Türkiye's internet infrastructure was extensive but had limited speed. In 2011, Türkiye's Information and Communication Technologies Authority (ICTA) decided not to regulate the incumbent fixed operator, Turk Telekom, for a period of five years or until fiber subscribers represented less than 25% of total fixed broadband subscribers. Turk Telekom agreed to offer wholesale access (Re-sale and Bit-Stream Access) to its fiber network under equal and non-discriminatory conditions. Additionally, to protect the rights of competing ISPs, ICTA mandated

Figure 1: **Optical Fiber Cable Roll-out in Türkiye**




**Note**: This figure depicts the roll-out of optical fiber cables and its breakdown between the backbone of the network and peripheral fibers laid to reach farther locations (fiber to the home) across Turkish provinces during the period 2012-2019. It is based on data obtained from the ICT Authority in Türkiye. Over 2012-2019, the length of optical fiber 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%. See Figure A2 for a map of changes in fiber length across provinces.

that Turk Telekom continue providing wholesale services on its existing network in areas undergoing fiber transition.

This deregulation was accompanied by a key implementation detail: private ISPs were permitted to utilize optical fiber cables laid out by BOTAS, the sole natural gas and oil distributor in the country. Installed alongside the country's pipeline network prior to the reform, this infrastructure effectively served as the backbone for nationwide fiber expansion.

Between 2011 and 2019, the length of optical fiber cables in Türkiye nearly doubled (see Figure 1). The total length of cables reached 390,800 kilometers – equivalent to about 0.48 kilometers per square kilometer of land area – with investment directed primarily toward fiber-to-the-home (FTTH) rather than the backbone. Adoption was substantial: the number of fiber subscribers increased fivefold during this period, and fiber's share in total number of fixed broadband subscribers rose steadily (see Figure A1). As of 2020, fiber internet accounted for 23.9% of all fixed broadband subscribers in Türkiye, converging to the OECD average of 30.6%.

## 2.2 Why Fiber Infrastructure Matters

Fiber-optic internet is often considered superior to other broadband technologies like DSL (Digital Subscriber Line) or cable internet for several reasons. First, fiber 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 fiber-optic internet ideal for bandwidth-intensive activities like high-definition streaming and large file transfers. Second, fiber-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, fiber-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, fiber 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, fiber provides a more stable and predictable performance.[2]

## 2.3 Data Sources

We bring together several complementary datasets:

**Firm-to-Firm Trade and Balance Sheets.** Administrative data from the Ministry of Industry and Technology (MoIT) covering all formal firms from 2011–2019. This includes VAT-based records of domestic firm-to-firm trade, income statements with gross sales, employment, and wages, and registry data with location (province) and 4-digit NACE industry code – the standard industry classification in the European Union.[3] The full sample covers about 1.5 million firms in 494 industries, forming over 27 million links. We focus on manufacturing: about 265,000 firms, 226 4-digit industries, and 4 million links. On average, a firm has 9.3 suppliers, with 4.7 per buyer–origin province.

**ICT Usage Survey.** An annual survey by the Turkish Statistical Institute (TUIK), conducted since 2005, covering roughly 10,000 firms each year. It includes all firms with more than 250 employees and a representative sample of smaller firms. We use these data to measure firm-level adoption of high-speed internet, defined as reported connections exceeding 100 Mbps.

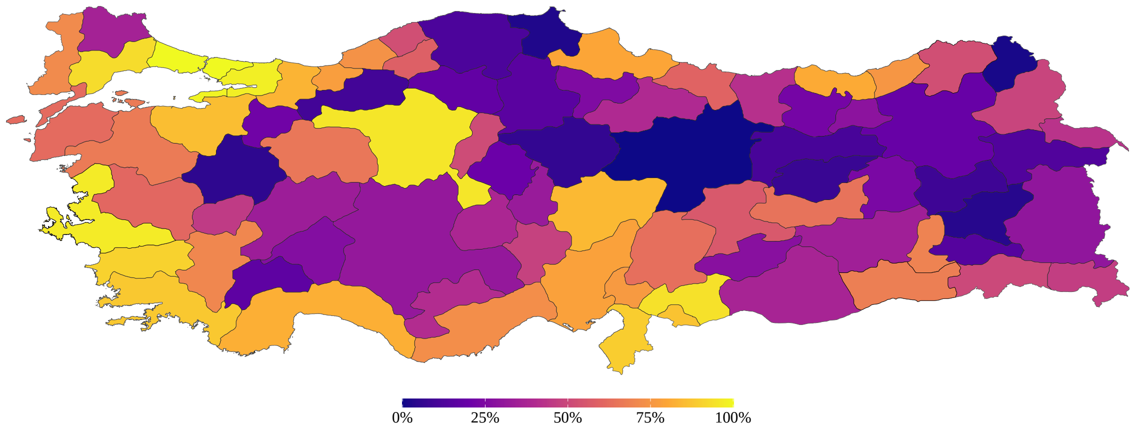
**Internet Infrastructure and Subscriptions.** Province-year data from ICTA on fiber cable rollout, broadband subscriptions (fiber, DSL, cable, mobile), and the GIS backbone network. For 2016–2019 we add Ookla's province-level upload and download speeds, which allow us to study not only adoption but also reliability of internet access.

**Regional and Geographic Controls.** Province- and district-level economic outcomes (GDP, population, employment, urbanization) from TUIK. Road network GIS data (Cosar et al. 2021) provide measures of inter-province travel time, which serve as controls in some specifications.

[2] According to a survey conducted by *Which* magazine among 3,000 participants, 63% reported faster speeds, 49% experienced fewer connection dropouts, and 39% reported fewer prolonged outages.

[3] VAT data record transactions above a non-binding threshold of 5,000 TL ($\approx$ US$ 840 in 2019).

Figure 2: **Change in Fiber Intensity**




**Note**: The map displays the change in fibre intensity between 2012 and 2019 across the provinces. Colors encode the percentile rank of the change, with lighter shades indicating larger increases in fibre intensity.

**BOTAS Pipeline Network.** Digitized 2011 GIS map of the BOTAS oil and gas pipelines, which also carry optical fiber cables. We use population-weighted district distances to the BOTAS network to construct instruments for province-level fiber intensity (details explained later).

### 2.4 Measuring Fiber Connectivity Across Provinces

We measure province-level fiber deployment intensity using the density of installed optical fiber cable. For province $d$ in year $t$, it is defined as

$$I_{dt} = 1 + \frac{L_{dt}}{A_d},$$

where $L_{dt}$ denotes the length of optical fiber cables (in kilometers) in province $d$ and $A_d$ is its land area (in km$^2$). Normalizing by area converts deployment into a spatial density measure, i.e., how much fiber infrastructure is available per unit of space, which is the natural scaling for geographically dispersed network capital. Figure 2 shows the change in intensity across provinces between 2012 and 2019. The median province experienced a 68% increase, with Istanbul at the high end (177%) and Kütahya at the low end (31%).

### 2.5 Construction of the Instrument

Our instrumental variable strategy draws on the historical geography of Türkiye's natural gas infrastructure. Long before broadband deregulation, the state-owned energy company BOTAS had laid optical fiber cables alongside its oil and gas pipelines. These cables were installed to support operational monitoring and control – enabling the company to detect leaks, monitor pressure, and coordinate system maintenance in real time – not to provide public internet access.

The 2011 reform, which allowed private internet service providers to lease and extend this infrastructure, transformed a pre-existing industrial communications network into the backbone of the country's fiber optic system. This institutional change meant that the spatial pattern of high-speed internet rollout was shaped, in large part, by the existing location of the BOTAS pipeline network. Because the pipelines had been routed decades earlier according to engineering and energy transport needs rather than areas of high economic activity, their configuration is plausibly unrelated to contemporary patterns of productivity or internet demand.

Importantly, the physical proximity of a province to the BOTAS network lowered the cost of expanding fiber-to-the-home (FTTH) connections once the market was opened to private investment. Laying new fiber requires expensive civil works such as trenching and duct installation. These costs increase steeply with distance from the existing backbone. Provinces whose population centers are located closer to the BOTAS network could, therefore, expand coverage at lower marginal cost and at a faster pace, while those located farther away faced higher costs and slower rollouts. In this way, proximity to the pre-existing pipeline network led to lower costs of FTTH investment and earlier adoption of high-speed broadband.

This spatial pattern is evident in Figure A3, which shows the BOTAS pipeline network overlaid on Türkiye's provincial boundaries. We calculate, for each district $m$, the minimum distance from its center to the BOTAS pipeline network, $Z_m$. Province-level distances are then computed as population-weighted averages of district distances in 2011:

$$\overline{\text{Distance}}_o = \sum_{m \in o} \frac{\text{Population}_{m,2011}}{\text{Population}_{o,2011}} \times Z_m. \tag{1}$$

This province-level measure captures how close the typical resident of province $o$ was, before deregulation, to the existing backbone network – and hence the relative cost of connecting to it once private ISPs were allowed entry. Figure A4 shows the spatial variation in the measure, highlighting how proximity to the BOTAS network differs markedly across regions of Türkiye.

In Table 1, we document that $\overline{\text{Distance}}_o$ is orthogonal to pre-reform province characteristics (2011), including GDP per capita, the share of adults with tertiary education, and the share of internet subscribers: the estimated coefficients are small and statistically insignificant. Therefore, the geography of the BOTAS pipeline network offers a natural experiment: it generated variation in infrastructure costs that was pre-determined and unrelated to existing economic conditions.

In the first stage of our instrumental variable approach, we estimate the relationship between fiber rollout and proximity to the BOTAS network at the province-year level. Specifically, we regress the observed length of optical fiber cables (in logs, normalized by area) on the interaction between provincial distance to the BOTAS network and year dummies for 2012–2018, controlling for province fixed effects, year fixed effects, and a rich set of time-varying province characteristics,

Table 1: **Distance to BOTAS Network and Initial Province Characteristics**

| Dependent Variable: IV: $\overline{\text{Distance}}_o$ | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Log Internet Penetration (2011) | 0.449 | | | | 0.458 |
| | (0.318) | | | | (0.462) |
| Log GDP per Capita (2011) | | 0.0424 | | | 0.00565 |
| | | (0.0297) | | | (0.00869) |
| Log Area | | | 0.00107 | | 0.00153*** |
| | | | (0.00104) | | (0.000527) |
| Tertiary Education Share (2011) | | | | 1.138* | -0.0616 |
| | | | | (0.628) | (0.628) |
| Observations | 81 | 81 | 81 | 81 | 81 |
| F stat. | 1.989 | 2.041 | 1.056 | 3.282 | 3.579 |

**Note**: Each observation pertains to province. Internet Penetration is the ratio of internet subscribers to population. Tertiary Education Share is the share of population with tertiary education. * 10%, ** 5%, *** 1% significance levels. Robust standard errors are reported in parentheses.

including initial per capita income of the province:

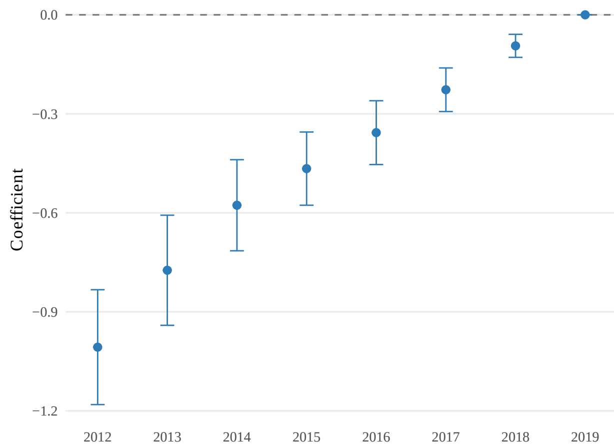
$$\ln I_{ot} = \alpha_o + \alpha_t + \sum_{\tau=2012}^{2018} \delta_\tau \left(\overline{\text{Distance}}_o \times D_\tau\right) + \mathbf{X}'_{ot}\gamma + \varepsilon_{ot}, \tag{2}$$

where $\overline{\text{Distance}}_o$ is defined in equation (1), and $D_\tau$ are year dummies for 2012-2018. The vector $\mathbf{X}_{ot}$ includes initial province characteristics such as logged GDP per capita, area, internet subscription rates, and the share of adults with tertiary education, each interacted with year dummies. Province and year fixed effects, $\alpha_o$ and $\alpha_t$, control for time-invariant provincial characteristics and aggregate shocks, respectively.[4]

Figure 3 presents the estimated coefficients on the interaction terms ($\overline{\text{Distance}}_o \times D_\tau$), with 2019 as the omitted year. All coefficients are negative, indicating that provinces farther from the BOTAS network had lower fiber intensity relative to the omitted year throughout the sample period. The coefficients are largest in magnitude in the early years of the rollout (2012–2013) and converge toward zero by 2018, indicating that the distance disadvantage attenuated over time. This pattern is consistent with the logic of the instrument: fiber deployment began near the BOTAS backbone, where expansion costs were lowest, and progressively extended to more distant provinces as the rollout matured. The joint significance of the interaction terms is strong, with a Kleibergen–Paap F-statistic exceeding 160, ruling out concerns about weak identification.

[4]Our first-stage specification is an "exposure design" in the spirit of Goldsmith-Pinkham, Sorkin, and Swift (2020): a time-invariant cross-sectional exposure is interacted with year dummies to generate plausibly exogenous time variation in fiber rollout. With province and year fixed effects, identification comes from differential changes in fiber deployment across provinces with different baseline proximity to the backbone. The key identifying restriction is that conditional on province and year effects and baseline province characteristics interacted with year dummies, baseline proximity does not proxy for other time-varying shocks that differentially affect fiber deployment (or our outcome variables) over 2012–2018.

Figure 3: Distance to BOTAS pipelines and First Stage Estimates




**Note**: This figure plots the coefficient estimates and the corresponding 95% confidence intervals obtained from first-stage regression of fiber connectivity on distance to BOTAS pipelines interacted with year dummies (see equation (2)). The distance of a province to BOTAS pipelines is constructed as the weighted average of the distances of districts within the province where district populations are used as weights. The Kleibergen–Paap $F$-statistic is 161.7 (standard errors clustered at the province level).

## 2.6 Empirical Validation: Internet Quality and Firm Connectivity

As a first pass, we examine whether the fiber rollout translates into observable improvements in internet quality, firm-level adoption of high-speed broadband, and early changes in firm behavior. This exercise also serves as an empirical validation of our identification strategy: if proximity to the BOTAS network indeed captures informative variation in the cost of expanding fiber infrastructure, then provinces predicted to experience greater fiber rollout should exhibit faster and more reliable internet connections, higher firm adoption rates, and behavioral shifts consistent with reduced communication frictions.

**Internet Quality.** We first examine whether provinces with higher fiber intensity experience measurable improvements in internet speed and reliability. Using Ookla's province-level data for 2016–2019, we estimate 2SLS regressions of average download and upload speeds, as well as their within-province variances, on fiber intensity while controlling for province and year fixed effects and interactions between baseline provincial characteristics and year dummies. Table A1 reports the results. A one percent increase in fiber connectivity is associated with about 2.4 and 4.8 percent higher average download and upload speeds, respectively, and a significant reduction in within-province variation in speeds. These findings indicate that fiber rollout translates into observable improvements in internet performance, supporting the validity of the instrument as a measure of broadband quality rather than physical infrastructure expansion.

**Firm-Level Adoption.** Next, we assess whether the availability of better internet infrastructure leads to greater adoption of high-speed internet by firms. Using the annual ICT Usage Survey, we estimate firm-level regressions of an indicator for connections exceeding 100 Mbps on provincial fiber intensity. The roll-out went hand-in-hand with high take-up by firms. Figure A6 shows that the share of firms with high-speed internet followed the trend in roll-out of fiber internet since 2011. To investigate how strongly adoption responded to the roll-out of fiber internet, we estimate the following equation for the 2012-2019 period:

$$\text{Adoption of High-Speed Internet}_{ft} = \gamma \ln I_{dt} + \alpha_d + \alpha_{st} + \text{Size}_{ft} + e_{ft} \tag{3}$$

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. The 2SLS estimate of the coefficient of interest is positive and statistically significant (see Table A2), implying that firms located in provinces with greater fiber rollout are more likely to adopt fast internet. The magnitude suggests that a 10 percent increase in fiber availability (equivalent to a 0.10 log-point rise) is associated with an almost 4 percentage-point increase in the probability of high-speed adoption. To gauge the economic relevance of this effect, note that the median province-level fiber intensity reached 0.31 (in logs) by 2019, with a 10*th*–90*th* percentile range of about 0.4 log-points—roughly a 50 percent difference in actual fiber coverage. This variation implies a 16 percentage-point higher adoption rate in the best-connected provinces. Given that the national share of firms with high-speed internet increased from virtually zero in 2011 to about 30 percent by 2019, the magnitude of our estimate suggests that differences in fiber rollout across provinces can explain a substantial share of the observed diffusion of high-speed internet among Turkish firms.

**Firm Behavior.** Finally, we link improvements in predicted fiber connectivity to firm outcomes using administrative data from the MoIT, which record all domestic firm-to-firm transactions. Table A3 reports 2SLS estimates of three outcomes: (i) the average geographic distance to suppliers, (ii) the share of input purchases from wholesalers, and (iii) total purchases from ICT-related industries. The results indicate that firms in provinces with greater fiber connectivity source inputs from suppliers located farther away, consistent with a decline in the costs of remote coordination and communication. The estimated coefficient implies that a 10 percent increase in fiber connectivity raises the average supplier distance by about 0.7 percent, suggesting that broadband expansion enables firms to integrate more distant suppliers into their production networks.

Firms also rely less on intermediaries: the share of purchases from wholesalers falls significantly in provinces with higher predicted fiber rollout. This reduction in intermediation points to more direct linkages between producers and end users, a pattern consistent with lower search and information frictions in input markets. Finally, we find that firms in better-connected provinces purchase more from ICT-related industries, indicating that improved broadband access stimulates complementary investment in digital technologies and services.

Taken together, these findings suggest that the rollout of high-speed internet not only improved communication infrastructure but also changed the way firms interact within domestic production networks. Enhanced broadband access appears to facilitate more complex, technology-intensive transactions and broader geographic reach, validating that our measure of fiber expansion captures substantive improvements in digital connectivity rather than nominal infrastructure growth. Collectively, these patterns indicate a broad reorganization of production networks that we document next, as the expansion of fast internet reshaped inter-provincial input sourcing and supplier diversification.

## 3 Fiber Connectivity and Input Sourcing

The previous section showed that broadband infrastructure expansion translated into faster and more reliable internet, greater firm-level adoption of high-speed connections, and measurable changes in firms' use of digital technologies. We now turn to our main question of how this expansion in digital connectivity affected the structure of production networks across space. In particular, we study whether firms reallocated and diversified their input sourcing toward better-connected regions following the rollout of fiber-optic internet.

**Measuring Bilateral Fiber Intensity.** To study the effects of improved internet access on inter-provincial trade in intermediate inputs, we extend our provincial measure of fiber intensity to the bilateral level. Many of the interactions that sustain modern production networks – negotiating contracts, exchanging technical specifications, resolving problems in real time, or sharing large digital files – depend on synchronous communication between firms. Such interactions require both parties to have sufficiently fast and stable internet connections to support smooth communication and real-time data transmission. When bandwidth is limited or connectivity unstable on either side, these exchanges become slower, more error-prone, and costlier to sustain. In practice, therefore, the effectiveness of digital communication between two firms is determined by the weaker of their two connections.

This technological complementarity motivates the construction of a bilateral measure of digital connectivity that reflects the "weakest-link" nature of communication. We define *bilateral fiber connectivity* between supplier province $o$ and buyer province $d$ in year $t$ as:

$$\ln \widehat{I}_{od,t} = \min\{\ln \widehat{I}_{ot}, \ln \widehat{I}_{dt}\}, \tag{4}$$

where $\widehat{I}_{ot}$ denotes the predicted value of fiber intensity for province $o$ in year $t$, obtained from the first-stage regression described in Section 2. Specifically, $\widehat{I}_{ot}$ represents the fitted value of the logarithm of optical fiber length per square kilometer, instrumented by the interaction of provincial distance to the BOTAS pipeline network with year dummies and controlling for province and year fixed effects as well as initial provincial characteristics interacted with annual dummies. Using predicted rather than observed fiber intensity isolates the exogenous component of broadband expansion that stems from historical and regulatory factors that are plausibly unrelated to contemporaneous economic conditions.

This bilateral measure provides a natural proxy for the effective level of digital connectivity between provinces. It captures the fact that the benefits of fast and reliable internet for inter-provincial transactions – such as improved communication quality, faster information flow, and enhanced coordination of production – arise only when both trading partners have adequate digital infrastructure. Figure A5 illustrates the evolution of this measure between 2012 and 2019, showing a noticeable rightward shift consistent with the rapid diffusion of fiber connectivity across provinces.

We next exploit variation in this bilateral measure of connectivity to examine how improved internet access shaped firms' sourcing decisions across Turkish provinces.

### 3.1 Empirical Strategy

To understand the effect of improved digital connectivity on firms' input sourcing decisions, we estimate how changes in bilateral fiber connectivity between provinces affect the intensity and composition of firm-to-firm trade links over time. Specifically, we relate measures of firms' input sourcing patterns across origin provinces to our bilateral measure of predicted fiber connectivity described above.

Our baseline empirical specification takes the following form:

$$\ln y_{ob,t} = \iota \ln \widehat{I}_{od,t} + \delta_{bt} + \delta_{ot} + \delta_{od} + \varepsilon_{ob,t}, \tag{5}$$

where $b$ indexes buyer firms, $o$ indexes supplier (origin) provinces, $d$ denotes the (destination) province in which buyer $b$ is located, and $t$ indexes years. The dependent variable, $y_{ob,t}$, captures different dimensions of the buyer-supplier relationship between firm $b$ and suppliers located in province $o$, as we explain below.

The specification includes an extensive set of fixed effects. Origin–destination fixed effects ($\delta_{od}$) capture all time-invariant determinants of bilateral trade, including geography, historical relationships, and transport costs. Origin–year fixed effects ($\delta_{ot}$) absorb any time-varying shock specific to a supplier province — such as new energy infrastructure investments, industrial policy interventions, or regional development programs. Buyer–year fixed effects ($\delta_{bt}$), which subsume destination–year effects, similarly absorb time-varying shocks to the buyer's province, including local demand conditions, regulatory changes, or public investment. The only remaining threat to identification is therefore a shock that varies at the origin–destination–year level and is correlated with bilateral fiber connectivity. Among such confounders, changes in bilateral transport infrastructure — such as road upgrades between specific province pairs — are the most natural candidate. We address this directly in one of our robustness checks, Table A11 by interacting baseline bilateral travel times with year dummies. Our coefficient on fiber connectivity remains stable in this specification. As a further check, we apply the same specification using a placebo measure of bilateral connectivity based on cable-TV subscription rates (Table A10). Cable-TV networks share similar physical infrastructure with fiber networks but were not used for business-to-business communication in Türkiye during this period. The null results reinforce the interpretation that our estimates capture the effect of high-speed internet rather than broader infrastructure corridor effects.

Therefore, the identifying assumption in specification (5) is that changes in the rollout of fiber infrastructure affect the evolution of buyer–supplier relationships across provinces only through their impact on internet connectivity. The rich set of fixed effects ensures that the estimated coefficient $\iota$ captures within-pair temporal changes in trade relationships that are associated with improvements in bilateral connectivity, net of local and aggregate shocks. After presenting the baseline results, we subject them to a battery of robustness checks. Standard errors are clustered two ways by origin and destination province to allow for arbitrary correlation of shocks within trading regions over time.

### 3.2 Reduced-Form Estimation and Baseline Results

**Reallocation Across Origins.** We begin by examining whether improved connectivity leads firms to reallocate their input purchases toward better-connected provinces. For each buyer firm $b$ and year $t$, we compute the share of total input spending that comes from suppliers located in each origin province $o$:

$$\text{Cost Share}_{ob,t} = \frac{\sum_{s \in o} \text{Purchases}_{sb,t}}{\sum_{s'} \text{Purchases}_{s'b,t}},$$

where $\text{Purchases}_{sb,t}$ denotes the total value of purchases made by buyer $b$ from supplier $s$ in year $t$. This measure captures how much of a buyer's material inputs are sourced from a given origin province.

We then estimate equation (5) using $\ln(\text{Cost Share}_{ob,t})$ as the dependent variable. The 2SLS results, reported in column (1) of the lower panel of Table 2, show that the coefficient on bilateral fiber connectivity is positive and highly significant. The estimated elasticity of 0.43 implies that a 10 percent increase in bilateral connectivity raises the share of inputs sourced from that province by approximately 4.3 percent. The magnitude is economically meaningful when compared to the observed variation in predicted broadband connectivity and sourcing patterns. The mean share of inputs that a buyer sources from a given province is about 10 percent, corresponding to a mean log value of –2.3 in our data. The estimated elasticity, therefore, suggests that a move from the 25*th* to the 75*th* percentile of connectivity – an increase of about 0.4 log points or 50 percent in broadband availability – raises the share of inputs sourced from that origin by approximately 18 percent, or nearly 2 percentage points relative to the mean. These magnitudes indicate that the rollout of broadband connectivity explains a quantitatively important portion of the reallocation of intermediate trade toward better-connected provinces, consistent with the interpretation that improved internet access substantially reduced the costs of coordination and communication across regions. Overall, our results indicate that firms systematically reallocate their purchases toward suppliers located in regions with better digital infrastructure, consistent with lower communication costs and smoother coordination.

Table 2: **Reallocation and Diversification in Input Sourcing**

| Dependent Variable: | Cost Share (1) | No. Suppliers (2) | Cost Share HHI (3) | New Connections (4) |
|---|---|---|---|---|
| **Panel A: OLS** | | | | |
| Fiber Connectivity | 0.510*** | 0.325*** | -0.107*** | 0.0632*** |
| | (0.0549) | (0.0297) | (0.0136) | (0.0196) |
| **Panel B: IV** | | | | |
| Fiber Connectivity | 0.427*** | 0.261*** | -0.0846*** | 0.0489*** |
| | (0.0469) | (0.0268) | (0.0118) | (0.0175) |
| **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. *Number of Suppliers* counts the distinct suppliers of the buyer firm located in a given origin. *Cost Share HHI* is the Herfindahl–Hirschman Index of purchase concentration across suppliers within the origin province. *New Connections* counts newly established supplier relationships relative to the previous year. Panel B uses Fiber Connectivity constructed using the instrument's predicted values, as defined in (4). * 10%, ** 5%, *** 1% significance levels. Standard errors are two-way clustered by origin and destination provinces.

**Diversification Within Origins.** Next, we investigate whether improved connectivity also affects the structure of firms' supplier portfolios within provinces. To do so, we estimate equation (5) using three alternative dependent variables that capture different aspects of within-origin diversification.

First, we consider the *number of suppliers* in province $o$ that firm $b$ purchases from in year $t$. Column (2) of Table 2 shows that fiber connectivity between provinces $o$ and $d$ has a positive and statistically significant effect on the number of suppliers. The coefficient of 0.26 on the number of suppliers indicates that a 10 percent increase in bilateral fiber connectivity raises the number of suppliers a buyer trades with in a given origin by about 2.6 percent. This implies, a move from the 10*th* to the 90*th* percentile of predicted bilateral connectivity (an increase of $\approx 1.42$ log points) raises the expected number of suppliers by about 45%, or almost one additional supplier at the mean (1.8 suppliers).

Second, we examine the *concentration* of input purchases across suppliers within a province. For each buyer–origin–year combination, we compute the Herfindahl–Hirschman Index (HHI) of input purchase concentration:

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

Column (3) of Table 2 reports a negative and statistically significant coefficient of –0.08, indicating that improved connectivity reduces the concentration of purchases within source provinces. Moving from the 10*th* to the 90*th* percentile of predicted connectivity lowers the HHI by about 11 percent, from an average of 0.72 to roughly 0.64. This decline suggests that firms spread their purchases more evenly across suppliers and become less dependent on dominant partners as internet infrastructure improves.

Finally, we explore whether improved connectivity facilitates the formation of *new supplier relationships*. Column (4) shows that the number of new supplier links that a firm establishes with suppliers in province $o$ rises with connectivity, with an elasticity of 0.05. These results together indicate that improved internet access not only shifts the geography of input sourcing toward better-connected regions but also promotes greater diversification and relationship formation within those regions.

**Robustness Checks.** We subject our results to a battery of robustness checks to ensure that the estimated effects of fiber connectivity on firms' input sourcing patterns are not driven by measurement choices, omitted variables, or specific features of the sample.

*Alternative Measure of Connectivity.* Our baseline specification defines bilateral fiber connectivity as the minimum of predicted fiber intensity between provinces, reflecting the "weakest-link" nature of synchronous communication. To test the sensitivity of our results to this definition, we construct an alternative measure based on the *similarity* of digital infrastructure between trading partners:

$$\ln \widetilde{I}_{odt} = -\left|\ln \widehat{I}_{ot} - \ln \widehat{I}_{dt}\right|.$$

This measure captures how aligned the two provinces are in terms of broadband development and is higher when both have similar levels of connectivity. Table A4 shows that the estimated coefficients retain their signs and statistical significance relative to the baseline specification when using this alternative measure of fiber intensity. These results confirm that the observed relationship between broadband expansion and firm-to-firm trade is robust to alternative representations of internet connectivity.

*Additional Controls.* We next include additional bilateral controls to rule out potential confounding channels. First, we add the absolute difference in GDP per capita between origin and destination provinces to capture the fact that economically similar regions may trade more intensively even absent broadband improvements. Second, we introduce a measure of 3G/4G mobile connectivity, which proxies for access to mobile broadband services that could complement or substitute for fixed-line fiber networks. As shown in Table A5, including these controls leaves the main coefficients on fiber connectivity virtually unchanged, confirming that the estimated effects are not driven by differences in regional development or concurrent expansions in mobile networks.

*Exploiting the Source of Variation in Connectivity.* Our bilateral connectivity measure is dominated by the trade partner whose fiber connectivity is of lower quality than the other. Therefore, we expect that improvements in fiber 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.

*Including Non-Manufacturing Suppliers.* Our baseline analysis focuses on manufacturing-to-manufacturing network, where supplier relationships are well-documented and relatively homogeneous. To test the external validity of our results, we expand the sample to include all supplier firms, including those in services, trade, and other sectors. The results, reported in Table A7, remain robust in both sign and significance, implying that improved broadband connectivity influences firm-to-firm transactions across a wide range of industries, not just manufacturing.

*Excluding Multi-Region Firms.* Because firms with establishments in multiple provinces might establish internal supplier networks or reallocate purchasing across their own affiliates, their inclusion could bias the estimated effects of external connectivity. To check this, we re-estimate the baseline specification excluding buyers and suppliers that operate in multiple provinces. Table A8 shows that the coefficient estimates remain qualitatively similar to the baseline sample, indicating that the results are not driven by the presence of multi-region firms.

*Long Differences Between 2011 and 2019.* Next, we estimate a long-differences specification comparing cumulative changes in input sourcing between 2011 and 2019 to cumulative changes in predicted fiber connectivity. This approach mitigates concerns about serial correlation in annual data and captures long-run effects of broadband expansion. The results, presented in Table A9, confirm that the provinces experiencing larger improvements in connectivity also experienced greater reallocation and diversification of input purchases, consistent with the panel estimates.

*Cable TV Rollout as Placebo Test.* To confirm that our results capture the effects of broadband communication rather than broader infrastructure deployment, we conduct a placebo test replacing fiber connectivity with the minimum of cable-TV subscription rates at origin and destination. While cable TV networks share similar physical infrastructure (coaxial and fiber trunk lines), cable-TV subscriptions in Türkiye during this period primarily served residential entertainment. Unlike fiber-optic internet subscriptions, which were widely adopted by firms for business-to-business communication and data exchange, cable-TV take-up among businesses was negligible. As expected, the coefficients on this placebo measure are statistically insignificant across all outcomes (Table A10), confirming that the observed effects stem from internet-enabled business communication rather than from physical network expansion per se.

*Interactions Between Travel Time and Year Dummies.* A potential concern is that improvements in fiber connectivity may coincide with changes in transportation infrastructure, which could also affect inter-provincial trade. To address this, we interact inter-province travel times (from the GIS-based road network data of Coşar et al. (2022)) with year dummies, thereby controlling for time-varying reductions in physical trade costs. This controls flexibly for the possibility that province pairs that are initially more or less remote follow different trade trajectories over time, which could confound the estimated effects of fiber connectivity. Table A11 shows that our coefficients on fiber connectivity remain stable, suggesting that the observed effects are not confounded by concurrent improvements in road transport or logistics.

*Excluding Origin-Destination Fixed Effects.* We further assess the role of transportation frictions by estimating the baseline specification without origin–destination fixed effects, which allows the bilateral travel-time variable to be identified. The results, reported in Table A12, indicate that travel time is strongly and negatively associated with firm-to-firm trade intensity across provinces. Longer travel times significantly reduce the cost share and number of suppliers from a given origin and increase purchase concentration, as reflected in the positive coefficient on the HHI. Importantly, even when controlling for travel time directly, the estimated coefficients on fiber connectivity remain similar to the baseline estimates, reinforcing that improvements in digital infrastructure exert an independent effect on trade link formation beyond what can be explained by transportation costs.

*Bilateral Instrument.* Our baseline specification uses predicted values from province-level first-stage regressions to construct bilateral fiber connectivity, which is a generated-regressor approach rather than classical 2SLS. To verify that the nonlinear mapping from province-level instruments to the bilateral treatment does not distort our estimates, we also estimate a specification that directly instruments bilateral connectivity using an explicit bilateral function of the underlying distance measures. Specifically, we construct $\overline{\text{Distance}}_{od} = \max\left\{\overline{\text{Distance}}_o, \overline{\text{Distance}}_d\right\}$, where $\overline{\text{Distance}}_o$ is defined in equation (1). Since the province farther from the backbone determines the binding constraint on bilateral connectivity, this instrument mirrors the weakest-link logic of our treatment measure. Because $\overline{\text{Distance}}_{od}$ is time-invariant, we interact it with year dummies and use these interactions as instruments for bilateral fiber connectivity in a standard 2SLS framework. Table A13 presents the results. The first stage is strong, with a high KP-statistic, and the estimated coefficients on bilateral fiber connectivity are quantitatively similar to those obtained under our baseline approach across all four outcomes, confirming that our main findings are not sensitive to the method used to construct the bilateral instrument.

*Event-Study Evidence.* Finally, we conduct an event-study analysis of the effects of improvements in predicted fiber connectivity. Our setting departs from a canonical event-study framework in two respects. First, our treatment variable is continuous. Second, even after discretization, treatment timing differs across units because province pairs adopted fiber infrastructure at different points in time. To address the first issue, we convert the continuous treatment into a binary status variable. Specifically, we calculate the median of predicted fiber intensity across all provinces and years and define the year of "treatment" for a province pair as the first year in which both provinces record predicted fiber intensity above the full sample median. The resulting treatment indicator therefore captures the transition from below- to above-median connectivity, while naturally generating staggered adoption across province pairs. We then estimate dynamic treatment effects using the approach proposed by Callaway and Sant'Anna (2021).[5] Figure A7 plots the estimated dynamic effects for the four main outcomes – cost share, number of suppli-

[5]There are several examples in the applied literature where continuous treatment variables are discretized for event-study analysis. Callaway et al. (2024) discuss constructing discrete treatment bins, with the number and cutoffs chosen to suit the empirical context. Danzer et al. (2024) employ the sample median to define above- and below-threshold regions, and Akerman et al. (2015) define an "event" as the year of the largest increase in broadband availability to convert a continuous rollout measure into a binary treatment. Our implementation follows these precedents: the median threshold reflects a meaningful structural break in fiber availability across provinces while maintaining sufficient balance across treated and untreated groups.

ers, purchase concentration (HHI), and new connections—based on this binary treatment variable constructed from the predicted fiber measure. The estimates show no significant pre-treatment trends and a clear post-treatment increase in sourcing and diversification outcomes, consistent with the baseline 2SLS findings.

Across all these exercises—alternative measures, additional controls, heterogeneous effects, and alternative samples—the estimated effects of fiber connectivity on firms' sourcing behavior remain positive, significant, and economically meaningful. The results are robust to accounting for differences in regional income, mobile connectivity, transportation infrastructure, and firm scope. Placebo tests yield null results, and both the event-study and long-difference specifications confirm the persistence and timing of the effects. Together, these findings provide strong evidence that the rollout of high-speed internet, rather than correlated economic trends, drove the reorganization of production networks across Türkiye over the sample period.

### 3.3 Discussion of Empirical Results

Our empirical analysis reveals two robust patterns. First, firms allocate a larger share of their expenditures to suppliers located in provinces with better access to high-speed internet. Second, conditional on sourcing from such provinces, firms diversify more evenly across suppliers. How should we interpret these findings?

We argue that high-speed internet reduces two distinct frictions that shape production networks. The first is what we call *communication costs*—the ongoing costs of coordinating with suppliers. Managing production schedules, exchanging design specifications, monitoring quality, and resolving problems require the transfer of large digital files, integration with cloud-based enterprise resource planning systems, or frequent high-resolution video conferences. In provinces with weak internet connectivity, these activities are slow, disrupted, and error-prone. By contrast, provinces with robust digital infrastructure enable smoother, more reliable interactions, encouraging firms to allocate a greater share of their purchases toward such suppliers.

The second friction is *information acquisition costs*—the costs of learning about potential suppliers. Firms increasingly rely on digital platforms to identify potential partners, to exchange technical specifications, and to compare offers. For example, virtual plant tours allow firms to verify production capacity without costly travel, and high-definition video conferencing makes it possible to hold detailed negotiations with multiple candidates in quick succession. Large digital files such as prototypes, CAD drawings, and compliance documents can be transmitted and assessed more easily when bandwidth is abundant and reliable. These features make it feasible for firms to expand the pool of suppliers they consider and to distribute orders more evenly across them.[6]

If this interpretation is correct, we would expect the *information acquisition* channel to matter

[6] It is important to distinguish between acquiring precise information about potential suppliers and simply becoming aware of their existence. The former depends on access to high-speed fiber internet, whereas the latter is primarily linked to broadband connectivity (Malgouyres et al., 2021).

more in industries where buyers face greater uncertainty about supplier productivity. To test this hypothesis, we disaggregate our data by supplying industry and augment the baseline specification to the buyer–origin–supplying industry–year level. We then interact bilateral fiber connectivity with an indicator for industries characterized by high productivity dispersion. The indicator, denoted *High Prod. Dispersion*, equals one for 2-digit NACE industries whose coefficient of variation of the sales-to-employment ratio in 2011 exceeds the cross-industry mean. Table 3 reports results from estimating the following augmented specification:

$$\ln y^k_{ob,t} = \iota_1 \ln I_{od,t} + \iota_2 \ln I_{od,t} \times \mathbf{1}[\text{High Prod. Dispersion}^k_o] + \delta_{bt} + \delta^k_{ot} + \delta_{od} + \varepsilon^k_{ob,t}, \tag{6}$$

where $k$ indexes the 2-digit NACE industry of the supplier, and $\delta^k_{ot}$ denotes origin-industry-year fixed effects.

Results in Table 3 confirm this prediction: the information acquisition effects are significantly stronger in high-dispersion industries, consistent with fiber expansion relaxing information frictions where supplier quality is most uncertain. Specifically, the positive and significant coefficients on the interaction term in column (2) indicate that fiber connectivity leads to a greater increase in the number of suppliers in industries with high productivity dispersion. Similarly, the negative and significant coefficient in column (3) shows that improved connectivity reduces purchase concentration more sharply in these industries. By contrast, the communication friction effect, captured by the cost share regression in column (1), is not differentially affected by industry-level productivity dispersion, as indicated by the non-significant coefficient on the interaction term. This pattern suggests that the communication cost channel operates similarly across industries regardless of supplier heterogeneity, while the information acquisition channel is more salient where ex ante uncertainty about supplier quality is greater.

Together, lower communication and information acquisition costs provide a unified explanation for the spatial reorganization of production networks following fiber expansion. We formalize these mechanisms in the model developed in the next section.

# 4 Theoretical Framework

With the empirical results in hand, we now lay out a general equilibrium framework of trade between provinces in which firms, facing imperfect information, make heterogeneous input-sourcing decisions. The framework serves two main purposes. First, it enables us to quantify the aggregate implications of improved access to high-speed internet that arise through the changes in sourcing patterns documented above. Because our empirical strategy exploits variation across province pairs and over time, the estimates identify relative shifts in sourcing but do not, on their own, deliver aggregate effects. Embedding these mechanisms in a model allows us to bridge this gap. Second, the framework provides a way to disentangle the underlying channels. The evidence indicates not only that firms reallocate purchases toward provinces with better internet access, but also that they diversify more evenly across suppliers located there.

Table 3: **Understanding the Channels: Communication vs Information Acquisition Costs**

| Dependent Variable: | Cost Share (1) | No. Suppliers (2) | Cost Share HHI (3) | New Connections (4) |
|---|---|---|---|---|
| Fiber Connectivity | 0.430*** | 0.264*** | -0.0861*** | 0.0514*** |
| | (0.0467) | (0.0270) | (0.0119) | (0.0177) |
| High Prod. Dispersion | 0.143 | 0.619*** | -0.369*** | 0.162 |
| × Fiber Connectivity | (0.552) | (0.271) | (0.142) | (0.125) |
| **Fixed Effects:** | | | | |
| Buyer×Year | ✓ | ✓ | ✓ | ✓ |
| Origin×Industry×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, a 2-digit NACE supplying industry, and a year. All variables are in natural logarithms. *Cost Share* is the fraction of purchases of a buyer from the origin province. *Number of Suppliers* counts the distinct suppliers of the buyer firm located in a given origin. *Cost Share HHI* is the Herfindahl–Hirschman Index of purchase concentration across suppliers within the origin province. *New Connections* counts newly established supplier relationships relative to the previous year. Fiber Connectivity is constructed using the instrument's predicted values, as defined in (4). * 10%, ** 5%, *** 1% significance levels. Standard errors are two-way clustered by origin and destination provinces.

The model captures these mechanisms by explicitly incorporating two distinct frictions: communication costs and information acquisition costs. Better internet access reduces communication costs, making coordination with chosen suppliers more reliable and less expensive. It also lowers information acquisition costs, enabling firms to identify and select lower-cost suppliers more effectively. Both channels reduce the effective input price index faced by firms and, in turn, lower their unit costs of production.

The welfare implications of these reductions, however, are not unambiguous. On the one hand, lower communication costs act as a direct reduction in the trade frictions between connected provinces, shifting the sourcing distribution toward those regions. On the other hand, lower information acquisition costs increase the likelihood that firms discover cheaper suppliers, which reduces sourcing costs but simultaneously intensifies competition. Provinces with relatively high initial costs may lose market share, while those with lower costs may gain. These reallocations feed back into equilibrium wages, potentially offsetting part of the initial input-cost advantage.

Thus, the aggregate welfare consequences of high-speed internet access depend on the interaction of these two forces. While both channels unambiguously reduce sourcing frictions at the micro level, the general equilibrium outcome for real income across provinces depends on how changes in input price indices interact with equilibrium adjustments in factor markets and trade flows.

## 4.1 Model Setup

The theoretical framework is a model of inter-provincial trade in Türkiye, building on Eaton and Kortum (2002), with endogenous firm-to-firm link formation as in Panigrahi (2021) and input sourcing under rational inattention as in Dasgupta and Mondria (2018). It features three key elements aligned with our empirical findings. First, in addition to the standard iceberg trade costs of shipping goods across provinces, firms face communication costs that capture the frictions of coordinating production, exchanging design specifications, and managing transactions with suppliers in different locations. These costs directly influence the attractiveness of sourcing from a given province. Second, firms are imperfectly informed about the characteristics of potential suppliers. To reduce this uncertainty, they allocate attention toward acquiring information. Attention is scarce and costly, but by allocating more of it, firms improve their knowledge of suppliers' cost efficiency and increase the likelihood of forming profitable links. Third, access to high-speed internet reduces both types of frictions. Better connectivity lowers the cost of synchronous communication (e.g., video calls, data sharing, and real-time coordination) and also decreases the cost of acquiring and processing information about potential suppliers (e.g., through digital directories, online platforms, and virtual due diligence).

The economy consists of multiple provinces, indexed by $o, d$ each endowed with a positive measure of firms and households. The model is static and aims to capture the long-run steady state of the economy.

**Manufacturing Production**

Firms in the manufacturing sector produce differentiated varieties and are heterogeneous in their idiosyncratic productivity. Each firm combines locally supplied labor with intermediate inputs sourced from potentially distinct sets of suppliers distributed across multiple provinces. Production involves performing a collection of tasks, each of which requires specific intermediate inputs, in addition to labor. The technology is assumed to exhibit constant returns to scale. In what follows, we suppress the identity of the buyer for simplicity of notation and describe the production process for an anonymous firm located in province $d$. The production function, defined over labor and a discrete number of tasks (indexed by $k$), takes the form:

$$y_d = \kappa z_d \ell_d^{1-\alpha} \left( \prod_{k=1}^{K} m_d(k)^{1/K} \right)^{\alpha},$$

where $y_d$ denotes the firm's output, $\ell_d$ is the amount of labor employed, $m_d(k)$ is the quantity of materials purchased to accomplish task $k$, $z$ is the firm's Hicks-neutral productivity, $\alpha$ is the materials share of costs, $K$ is the number of tasks required in production, and $\kappa = \alpha^{-\alpha} (1-\alpha)^{-(1-\alpha)}$ is a normalizing constant. This formulation of the production function in terms of the tasks performed by intermediate inputs is similar to that proposed by Eaton, Kortum, and Kramarz (2022).

As in Panigrahi (2021), for accomplishing any task, the outputs of potential suppliers are perfectly substitutable within tasks of a firm, and the outputs of the same supplier can be used across multiple tasks by the same firm. Furthermore, although the functional form across tasks may resemble a Cobb–Douglas aggregator, the elasticity of substitution across suppliers is not fixed at one. Instead, it is given by $\sigma \geq 1$, the value of which is based on how likely it is to find a lower-cost supplier as determined by the endogenous distribution of effective costs across provinces.

Firms are assumed to price competitively in the output market, setting prices at their unit cost of production. Given the production technology, cost minimization implies that a constant share of total costs is allocated to labor and to each of the $K$ tasks. Accordingly, the demand for labor and the demand for inputs required to complete each task are then determined by the local wage rate and by the effective costs of inputs the firm faces for each task. Since the production function is homogeneous of degree one in all inputs, the unit cost function is independent of the firm's output level and can be written in closed form as $c_d = \frac{1}{z_d} w_d^{1-\alpha} \left( \prod_{k=1}^{K} p_d(k)^{1/K} \right)^{\alpha}$, where $w$ denotes the local wage and $p_d(k)$ denotes the effective cost of accomplishing task $k$ for the firm.

Having derived the unit cost function, we now turn to the firm's choice of suppliers for each task. Because inputs from different suppliers are perfectly substitutable within a given task, the cost-minimizing firm selects the supplier that offers the lowest effective cost. Since tasks enter symmetrically into the production function, we suppress the task index $k$ for clarity. The effective cost of sourcing from supplier $s$ is given by $p_{sd} = c_s \tau_{od} / a_{sd}$ where $c_s$ denotes the supplier's marginal cost of production, $\tau_{od}$ denotes the iceberg trade cost of shipping the input from the supplier's province $o$ to the destination province $d$, and $a_{sd}$ denotes the task-specific match productivity. Firms are imperfectly informed about suppliers' marginal costs and match productivities, ex ante. They can allocate costly attention to learn more precisely about suppliers' marginal costs, while they rely on expected values for match-specific productivities. Supplier choice is, therefore, shaped jointly by trade costs, information acquisition costs, and expectations about match productivity which capture communication costs.

**Communication Costs**

Communication costs enter the framework through match-specific productivities, which capture how easily a buyer and supplier can coordinate to make their relationship function efficiently for a given task. We assume that these productivities are idiosyncratic and independent across suppliers and tasks, reflecting that a supplier who is a good fit for one task may not be equally effective for another within the same buyer's production process.

Formally, let $a_{sd}$ denote the match-specific productivity of supplier $s$, located in province $o$, for a buyer performing a given task in province $d$. We suppress the task index $k$ for clarity. These draws are assumed to follow a Fréchet distribution,

$$\mathbb{P}\left(a_{sd} \leq a\right) = \exp\left(-\phi_{od} a^{-\zeta}\right),$$

where $\zeta > 0$ governs the dispersion of draws and $\phi_{od}$ is a province-pair specific parameter. A higher value of $\phi_{od}$ corresponds to a higher expected quality of coordination between suppliers in $o$ and buyers in $d$, reflecting lower communication frictions.

### Information Acquisition Costs

We model firms as being rationally inattentive in the sense of Sims (2003). A firm begins with some prior knowledge about the distribution of potential suppliers but can choose to devote costly attention to learn more about their marginal costs. The decision of how much information to acquire is made optimally, trading off the expected reduction in production costs against the cost of processing additional information. In practice, this means that the firm's choice of suppliers can be represented directly as a choice over a probability distribution across available suppliers, conditional on prior knowledge. Formally, the firm's supplier choice problem for any given task is expressed as:

$$\log p_d = \min_{\Pi_d \in \Delta} \left\{ \zeta \, \mathbb{E}_\Theta \left[ \sum_s \pi_{sd} \ln \left( c_s \tau_{o(s)d} / a_s \right) \right] + \psi(\Pi_d, \Theta) \right\}, \quad \Delta = \left\{ \pi_{sd} \geq 0, \sum_s \pi_{sd} = 1 \right\} \tag{7}$$

where $p_d$ is the effective price of the task sourced by the firm located in $d$, $\Theta$ is the probability measure that denotes prior knowledge of the firm, $\psi(\Pi_d, \Theta)$ denotes the cost of acquiring information, $\Pi_d \equiv \{\pi_{sd}\}_s$ and $\pi_{sd}$ denotes the probability of choosing supplier $s$ for the task conditional on prior knowledge. Firms do not acquire information about match-specific productivities $a_{sd}$. Instead, they choose suppliers using the prior distribution of these draws. Because supplier choice ultimately depends on effective costs $p_{sd} = c_s \tau_{od} / a_{sd}$, the Fréchet shape parameter $\zeta$ governs how strongly latent coordination shocks translate into economically relevant cost differences across suppliers. When $\zeta$ is larger, a given difference in supplier costs implies a larger difference in the probability of being the lowest effective-cost supplier, so the value of information about costs is higher relative to a fixed amount of attention. We capture this by measuring expected sourcing payoffs in units scaled by $\zeta$, while the information cost remains measured in attention units.

As in the rational inattention literature (Sims, 2003; Matějka and McKay, 2015), we model information costs as the reduction in uncertainty achieved by conditioning sourcing decisions on information relative to an ex ante benchmark. Specifically, for a firm in province $d$, the information acquisition cost is

$$\psi(\Pi_d, \Theta) = \Omega(\mathbb{E}_\Theta[\Pi_d]) - \mathbb{E}_\Theta[\Omega(\Pi_d)],$$

where $\Theta$ denotes the firm's prior over payoff-relevant states and $\Omega(\cdot)$ measures the uncertainty associated with a given vector of choice probabilities. Economically, this cost is larger when the firm uses information to make its sourcing rule sharply state-contingent and highly selective, and smaller when sourcing remains broad and diffuse across suppliers.

Following Fosgerau, Melo, de Palma, and Shum (2020), we take $\Omega(\cdot)$ to be a generalized entropy function that extends standard Shannon entropy to allow for different levels of information resolution. For a firm located in province $d$, the uncertainty associated with a supplier-choice distribution $\Pi_d$ is

$$\Omega(\Pi_d) = -\sum_s \pi_{sd} \left[ \lambda_{o(s)d} \ln \pi_{sd} + (\lambda - \lambda_{o(s)d}) \ln \left( \sum_{s' \in o(s)} \pi_{s'd} \right) \right].$$

This nested entropy separates uncertainty at two margins: across suppliers within a given origin and across origins themselves. The parameter $\lambda > 0$ sets the overall scale of the entropy, while the origin-specific parameter $\lambda_{od} \in (0, \lambda)$ governs how information about origin $o$ is resolved within the buyer's sourcing problem. The ratio $\lambda_{od}/\lambda \in (0,1)$ serves as the nesting parameter: higher values place relatively more weight on sharply distinguishing individual suppliers within origin $o$ and less weight on broader origin-level information, while lower values support more diffuse sourcing across suppliers within that origin. At the extremes, $\lambda_{od}/\lambda \to 1$ corresponds to information being resolved at the supplier level, while $\lambda_{od}/\lambda \to 0$ corresponds to information being resolved only at the province level, so suppliers within an origin are effectively pooled.

**Role of High-Speed Internet Access**

To capture the effect of high-speed internet access on firms' input sourcing decisions, we introduce reduced-form relationships linking fiber internet connectivity to both communication costs and information acquisition costs. First, we specify that the scale parameter of the Fréchet distribution governing match-specific productivities also depends on connectivity with a constant elasticity,

$$\frac{\partial \ln \phi_{od}}{\partial \ln I_{od}} = \gamma, \quad \gamma > 0.$$

Higher connectivity thus raises $\phi_{od}$, increasing the expected quality of match-specific productivity draws and thereby reducing communication frictions between suppliers in $o$ and buyers in $d$. Second, we assume that the information acquisition parameter is given by

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

where $I_{od}$ denotes fiber connectivity between $o$ and $d$. This formulation implies that when $\eta > 0$, better internet access reduces the cost of acquiring information about potential suppliers.

Together, these two channels provide a tractable way to model how improvements in high-speed internet access influence both communication costs and the information acquisition costs entering firms' sourcing decisions.

**Local Non-Traded Sector**

To capture the role of activities outside the manufacturing sector, we introduce a local non-traded sector in each province. This sector encompasses non-manufacturing goods and services that are both produced and consumed locally, without crossing provincial borders. Production in the non-traded sector uses only local labor under constant returns to scale. Let $T_d$ denote the province-specific productivity level in this sector. Under perfect competition, firms in the non-traded sector set prices equal to marginal costs, which implies that the local price index for non-traded goods in province $d$ is given by $w_d / T_d$.

### Household Preferences

On the demand side, we assume that each household supplies one unit of labor inelastically to firms in its province and earns labor income equal to the local wage. Household consumption encompasses two broad categories: tradable manufacturing goods and locally produced non-traded services. Preferences are represented by a Cobb–Douglas aggregator across these two components,

$$u_d = \left(\frac{C_d}{\beta}\right)^{\beta} \left(\frac{C_d^*}{1-\beta}\right)^{1-\beta}$$

where $C_d$ denotes consumption of manufacturing goods, $C_d^*$ denotes consumption of local non-traded goods and services, and $\beta \in (0,1)$ is the expenditure share on manufacturing.

Preferences over manufacturing goods are modeled analogously to the production technology described above, with a structure that explicitly incorporates tasks. Let there be a finite number of tasks indexed by $k$. Consumption of manufacturing goods is then given by

$$C_d = \prod_{k=1}^{K} \left(\sum_s q_{sd}(k)\right)^{1/K}$$

where $q_{sd}(k)$ is the quantity of goods sourced from supplier $s$ to accomplish task $k$. This specification implies that utility takes the form of a symmetric Cobb–Douglas across tasks, with perfect substitutability among suppliers within each task, consistent with the firm-side technology.

As in production, households are imperfectly informed about the full set of available suppliers ex ante. They face information frictions in learning about suppliers' characteristics and therefore allocate limited attention optimally across tasks. In line with the rational inattention framework, households choose a vector of conditional choice probabilities over suppliers, balancing the expected gains from more accurate supplier selection against the cost of acquiring and processing information.

## 4.2 Equilibrium

With the model primitives established, we now characterize the equilibrium of the model economy. The exposition proceeds in several steps. First, we derive the supplier choice probabilities that govern firm-to-firm trade. These probabilities formalize how trade costs, information frictions, and communication costs jointly determine the allocation of sourcing across potential suppliers, and thereby shape the structure of production networks.

Next, we present the market-clearing conditions for goods produced by manufacturing firms, consistent with the derived supplier choice probabilities. In the non-tradable sector, market clearing implies that local prices are pinned down by local productivity and wages. In each province's labor market, equilibrium requires that the inelastic local supply of labor equals total demand from firms operating in that province.

We then define household welfare as a function of labor income and the expected prices of both tradable goods produced by manufacturing firms and non-tradable goods and services. This welfare measure provides the basis for evaluating counterfactual equilibria.

Finally, we outline how the framework is used to quantify the aggregate welfare gains from improved access to fiber-optic internet. The analysis isolates the distinct contributions of reduced communication costs and information acquisition costs, thereby disentangling their respective roles in shaping equilibrium sourcing patterns and aggregate welfare outcomes.

#### Supplier Choice Probabilities and Firm-to-Firm Trade

We characterize firm-to-firm trade through the posterior probabilities of supplier selection, which arise as the solution to the rationally inattentive firm's input sourcing problem. These probabilities represent the equilibrium distribution over potential suppliers that minimizes the expected cost of inputs subject to an information acquisition constraint. They are conditioned on the marginal costs of potential suppliers but are taken in expectation over match-specific productivities, which remain unknown to the buyer at the time of supplier choice. Because match-specific productivities are assumed to be independent and identically distributed across tasks within a firm, the probability that a given supplier $s$ is selected for one task is identical to the probability of being selected for any other task.

Given the firm's input sourcing problem described in equation (7) and the specified functional forms for communication costs and information acquisition costs, we can derive a closed-form characterization of these posterior probabilities. The resulting expression defines, for any firm located in province $d$, the conditional probability of choosing supplier $s$ from the set of potential suppliers, as summarized in the following proposition. This characterization provides the microeconomic foundation for the formation of inter-firm links in equilibrium, which in turn determines the structure of the production network.

**Proposition 1.** *The probability with which any firm located in $d$ selects supplier $s$ for any given task, conditional on its marginal cost of production being $c_s$, is*

$$\begin{aligned} \pi_{sd} &= \pi_{sd|o(s)} \times \pi_{o(s)d}, \\ \pi_{sd|o(s)} &= \frac{\bar{\pi}_{sd|o(s)} c_s^{-\zeta/\lambda_{o(s)d}}}{\sum_{s' \in o(s)} \bar{\pi}_{s'd|o(s)} c_{s'}^{-\zeta/\lambda_{o(s)d}}}, \end{aligned} \tag{8}$$

$$\pi_{od} = \frac{\bar{\pi}_{od}\tau_{od}^{-\zeta/\lambda}\phi_{od}^{1/\lambda}\left(\sum_{s\in o}\bar{\pi}_{sd|o}c_s^{-\zeta/\lambda_{od}}\right)^{\lambda_{od}/\lambda}}{\sum_{o'}\bar{\pi}_{o'd}\tau_{o'd}^{-\zeta/\lambda}\phi_{o'd}^{1/\lambda}\left(\sum_{s'\in o'}\bar{\pi}_{s'd|o'}c_{s'}^{-\zeta/\lambda_{o'd}}\right)^{\lambda_{o'd}/\lambda}}, \tag{9}$$

*where* $\pi_{od}$ *denotes the probability of choosing a supplier from province o and* $\pi_{sd|o(s)}$ *the probability of choosing supplier s conditional on having chosen to source from the province where it is located,* $o(s)$.

See Appendix C.1 for the proof.

The posterior supplier choice probabilities stated above exhibit a structure closely related to that of a nested logit model, with the distinction that they are adjusted by the vector of prior probabilities, $\bar{\Pi}_d$. These priors reflect firms' ex ante beliefs about the relative attractiveness of sourcing from different provinces and suppliers before allocating attention or acquiring new information. As expected, higher prior probabilities translate into higher posterior probabilities, since firms that are initially perceived as more favorable are more likely to be selected once information is updated.

Equation (8) characterizes the probability of choosing a supplier within a given province, while equation (9) gives the probability of sourcing from that province. The inner probability reflects competition among individual suppliers and highlights that firms with lower marginal costs, after accounting for trade and information frictions, are more likely to be selected. The outer probability, in turn, captures the allocation of sourcing across provinces as a function of relative costs, information acquisition parameters, and communication efficiencies.

Importantly, the structure of these choice probabilities maps directly into observed firm-to-firm sales data, providing a transparent link between theory and empirics. This mapping facilitates the estimation of the model, as discussed in Section 5, and allows the comparative statics implied by the model to align closely with the reduced-form patterns documented in Section 3. The framework thus provides a coherent bridge between the micro-level mechanisms of supplier choice and the aggregate empirical relationships observed in the data.

**Role of High-Speed Internet Access**

The supplier choice probabilities together with our assumptions on the reduced-form relationships between the two cost channels and fibre connectivity shed light on how comparative statics arising from the model deliver the reduced-form results we presented in Section 3.

**Information Acquisition Costs and Sourcing Concentration:** Improvements in fibre connectivity shape the concentration of supplier sourcing within locations through their impact on information acquisition costs. In our framework, firms face a cost of acquiring information about potential suppliers in each origin province $o$. This cost determines how much attention firms optimally allocate across suppliers when deciding from whom to source.

Improved connectivity affects sourcing concentration through two opposing forces. First, higher $\zeta/\lambda_{od}$ sharpens the buyer's responsiveness to cost differences, which, holding the set of actively considered suppliers fixed, concentrates purchases on the lowest-cost suppliers. Second, lower information costs expand the set of suppliers about which it is profitable to acquire information, spreading attention across a broader pool and reducing concentration. As Dasgupta and Mondria (2018) show, when information costs are sufficiently high, so that buyers actively consider only a small subset of available suppliers, a decline in these costs generates diversification and HHI falls. Intuitively, the main effect of cheaper information in this regime is to bring previously ignored suppliers into the consideration set, rather than to sharpen discrimination among those already considered.

In our setting, the pre-fiber environment corresponds precisely to this high-cost regime: before the rollout of high-speed internet, the cost of acquiring information about distant suppliers was prohibitively high, limiting firms to a narrow set of known partners. The subsequent decline in information costs induced by fiber connectivity generates the diversification documented in Section 3.

**Communication Costs and Origin Sourcing Probabilities**: Beyond information acquisition, fiber connectivity directly affects firms' ability to coordinate and communicate with suppliers across space. In the model, this effect enters through a communication cost channel that reduces the effective cost of transacting with suppliers located in a given origin province $o$. Communication costs capture the difficulty of synchronous interaction which are particularly sensitive to digital infrastructure. Better internet quality $I_{od}$ lowers these costs, thereby increasing the attractiveness of sourcing from that province.

Fibre connectivity thus influences equilibrium sourcing probabilities not only by improving the information firms can obtain about suppliers (the information acquisition cost channel), but also by reducing the fixed and variable costs of interacting with them (the communication cost channel). The total effect of improved connectivity on bilateral sourcing from $o$ to $d$ can be obtained by differentiating $\ln \pi_{od}$ with respect to $\ln I_{od}$. Writing the origin-level numerator as $A_{od} = \bar{\pi}_{od} \tau_{od}^{-\zeta/\lambda} I_{od}^{\gamma/\lambda} S_{od}^{\lambda_{od}/\lambda}$, where $S_{od} = \sum_{s \in o} \bar{\pi}_{sd|o} c_s^{-\zeta/\lambda_{od}}$ is the within-origin inclusive value, a straightforward application of the chain rule yields

$$\frac{d \ln \pi_{od}}{d \ln I_{od}} = \frac{1 - \pi_{od}}{\lambda} \left( \gamma - \eta \lambda_{od}^2 \ln S_{od} - \zeta \eta \lambda_{od} \sum_{s \in o} \pi_{sd|o} \ln c_s \right). \tag{10}$$

The elasticity reflects three distinct forces, scaled by the inverse of the outer-nest entropy weight $1/\lambda$. The first term, $\gamma$, captures the direct communication channel: better connectivity raises the expected quality of match-specific productivity draws, making origin $o$ more attractive. The second and third terms arise from the information-acquisition channel. Because $1/\lambda_{od} = \bar{\eta} + \eta \ln I_{od}$, higher connectivity lowers $\lambda_{od}$, which affects the origin's attractiveness through two reinforcing forces: $-\eta \lambda_{od}^2 \ln S_{od}$ captures the variety effect (positive, since $\ln S_{od} > 0$, reflecting that sharper cost-sensitivity amplifies the value of deep supplier pools), and $-\zeta \eta \lambda_{od} \sum_s \pi_{sd|o} \ln c_s$ captures cost-selection reinforcement (also positive, since share-weighted log costs are typically negative).

The combined effect on $\pi_{od}$ is thus unambiguously positive when $\gamma, \eta > 0$: both the communication channel and the information-acquisition channel raise the sourcing probability. This comparative static implies that provinces experiencing larger improvements in fiber connectivity should exhibit stronger increases in incoming firm-to-firm links, consistent with the empirical patterns documented earlier.

**Market Clearing for Local Non-Tradables**

Consumers in province $d$ spend a constant share of their income $1-\beta$ on local non-tradables. This implies expenditure on local non-tradables is given by $(1-\beta)w_d L_d$. Demand for local non-tradables is then given by $(1-\beta)T_d L_d$.

**Labor Market Clearing** Since the local non-tradable sector prices competitively in the output market, its revenue from sales equals the cost of production. Given the technology of production, the cost of production is given by $w_d L_{d,NT}$ where $L_{d,NT}$ is the amount of labor employed in the non-tradable sector. Together, this implies that a share $(1-\beta)$ of workers in $d$ are employed in the non-tradable sector, $L_{d,NT} = (1-\beta)L_d$.

The remaining share $\beta$ of workers would then be employed in manufacturing. As per their production function, manufacturing firms in any province $d$ spend a constant share $1-\alpha$ of their inputs costs on labor. It then follows that the demand for labor by a manufacturing firm $b$ located in province $d$ can be expressed as $\ell(b) = \frac{1}{w_d}(1-\alpha)c_b y_d(b)$.

Putting these together, the total demand for labor by manufacturing firms in province $d$ is given by

$$\beta L_d = \sum_{b \in d} \ell_d(b).$$

Substituting respective values, the wage bill of manufacturing firms in $d$ can be expressed as:

$$\beta w_d L_d = (1-\alpha) \sum_{b \in d} c_b y_d(b).$$

**Market Clearing for Manufacturing Goods**

The total demand for manufacturing goods in province $d$ can be obtained by aggregating demand for intermediates by buyer firms and demand for final consumption by households, both potentially distributed across multiple provinces.

For a buyer firm $b$ located in province $d$, the expected share of tasks that supplier $s$ is selected for is given by $\pi_{sd}$ defined above. Since the production function of manufacturing firms features symmetric and Cobb-Douglas tasks, the expected share of expenditures on inputs of the buyer firm that is due to purchase of goods from firm $s$ is $\alpha\pi_{sd}$. Since manufacturing firms are assumed to price competitively in the output market, the expenditure on firm $s$ goods is $\alpha\pi_{sd}c_b y_d(b)$ where $y_d(b)$ denotes the output of buyer firm $b$. The demand for goods from $s$ is then $\alpha\pi_{sd}c_b y_d(b)/c_s$. Similarly, the demand for goods from $s$ for final consumption in $d$ is $\beta\pi_{sd}w_d L_d/c_s$. Aggregating across buyer firms and households spread across all provinces, the demand for firms $s$ goods is given by:

$$y_o(s) = \frac{1}{c_s} \sum_d \left( \sum_{b \in d} \alpha \pi_{sd} c_b y_d(b) + \sum_{h \in d} \beta \pi_{sd} w_d \right).$$

Noting that the supplier choice probabilities $\pi(s, d)$ are the same for all firms and households in $d$, this further simplifies to

$$c_s y_o(s) = \sum_d \left( \alpha \pi_{sd} \sum_{b \in d} c_b y_d(b) + \beta \pi_{sd} w_d L_d \right).$$

Expressing firm revenues in terms of their wage bill, we can rewrite this condition as:

$$w_o \ell_o(s) = \sum_d \pi_{sd} \beta w_d L_d.$$

Summing across all suppliers in $o$ then gives us the aggregate trade equilibrium condition:

$$w_o L_o = \sum_d \pi_{od} w_d L_d. \tag{11}$$

Equation (11) is the labor market clearing condition for each province. It equates the labor income (total wage bill of all firms) in $o$ to expenditure of goods produced at $o$, coming from final consumption and intermediate input usage in all provinces.

### 4.3 Welfare Analysis

The model lends itself naturally to welfare analysis. We define welfare as the indirect utility of households which itself depends on nominal wages, the price index for the non-tradable manufacturing sectors, and more crucially the expected price of manufacturing goods. The expected price of manufacturing goods for households coincides with the optimized value of the objective function for the firm's input sourcing problem. This is because the household's supplier choice problem is assumed to be analogous to that of the manufacturing firms. The following proposition characterizes the expected prices and hence the price index of manufacturing goods faced by households at each province, and hence welfare across provinces $\{V_d\}_d$.

**Proposition 2.** *The indirect utility is given by:*

$$V_d = \left( \frac{w_d}{P_d} \right)^{\beta} T_d^{1-\beta},$$

*where $P_d$ is given by:*

$$P_d = \exp\left(-\frac{\lambda}{\zeta}\mathbb{E}_\Theta\left[\ln\left(\sum_o \bar{\pi}_{od}\tau_{od}^{-\zeta/\lambda}\phi_{od}^{1/\lambda}\left(\sum_{s\in o}\bar{\pi}_{sd|o}c_s^{-\zeta/\lambda_{od}}\right)^{\lambda_{od}/\lambda}\right)\right]\right).$$

See Appendix C.2 for the proof.

**Gains from Trade** The model presented here departs from the class of trade models studied in Arkolakis et al. (2012) (ACR) in which welfare changes can be summarized by two sufficient statistics: the domestic expenditure share and a constant trade elasticity. In ACR-type frameworks, bilateral trade shares exhibit a constant elasticity of substitution with respect to trade costs, typically arising from CES preferences or Fréchet productivity draws with an invariant shape parameter. In our model, the trade elasticity is $\zeta/\lambda$, where $\lambda$ is the outer-nest entropy weight; but information and communication frictions endogenously alter the elasticity of substitution across suppliers. The derivative of sourcing probabilities with respect to fiber connectivity, shown in equation (10), depends on $\lambda_{od}$, which itself varies with the level of fiber connectivity $I_{od}$ and the distribution of supplier marginal costs within each origin. This feature breaks the constant elasticity structure that underpins ACR's sufficiency results. As a consequence, welfare changes in this environment cannot be summarized solely by changes in the domestic expenditure share; rather, they depend on how information and communication technologies reshape the effective distribution of sourcing elasticities and the composition of firm-to-firm links across origins. Furthermore, the model also departs from richer models featuring search and matching frictions but with constant elasticities studied in Arkolakis et al. (2025). Only in the special limiting case, when information acquisition costs are absent ($\eta = 0$) does the model nest the ACR structure as a particular case.

**Counterfactual Equilibria.** To construct the counterfactual equilibrium, we simulate the Turkish economy under the same exogenous fundamentals that prevailed in 2012, prior to the large-scale expansion of the optical fiber network. The only change is that the spatial distribution of fiber infrastructure is updated to reflect the average annual improvement observed between 2012 and 2019. To compute this counterfactual, we adopt the "exact hat algebra" methodology of Dekle et al. (2008), following its adaptations to settings with granular data in Dingel and Tintelnot (2020) and Panigrahi (2021). All equilibrium relationships are expressed in proportional changes relative to the baseline equilibrium, allowing the counterfactual to be solved without fully recomputing the level system.

A key distinction in our setting is that the elasticity of firm-to-firm trade with respect to supplier marginal costs varies with fiber connectivity, so the proportional neutrality that enables the full cancellation of primitives in standard hat algebra no longer applies. In addition, the welfare expression in Proposition 2 is written as an expectation with respect to buyers' prior information. In the quantitative implementation, however, the economy features a large number of buyers, and buyer-specific prior distortions are assumed to be idiosyncratic and mean-zero in the cross section. As a result, the aggregate prior entering the equilibrium allocation is well approximated by a uniform prior in the large-market limit. Under this aggregation, the prior-information expectation in Proposition 4 is equivalent to the deterministic price-index representation used in the counterfactual system. The hat-algebra exercise should therefore be interpreted as solving the counterfactual for the large-market aggregate economy, in which individual buyer-level prior heterogeneity washes out, while the non-constant sourcing elasticities induced by information frictions remain central for equilibrium and welfare. Accordingly, we explicitly recover the relevant primitives from the estimated structural model and use them to construct shocks representing changes in information acquisition and communication costs.[7]

We evaluate the welfare implications of improved connectivity by comparing household welfare in 2012 to its level in the counterfactual equilibrium where only the fiber network has expanded. This exercise isolates the component of welfare growth attributable to the rollout of optical fiber infrastructure, as transmitted through reductions in communication and information frictions, while holding constant all other economic fundamentals. The resulting equilibrium reflects both the direct reduction in intermediate input costs and the indirect general-equilibrium responses, including differential adjustments in wages across provinces.

# 5 Estimation and Calibration

The theoretical framework developed in the preceding section provides a foundation for quantifying the welfare effects of high-speed internet expansion. Doing so requires assigning values to the model's structural parameters and solving for counterfactual equilibria under alternative fiber connectivity scenarios. This section describes how we calibrate the model, estimate the key elasticities governing the response of sourcing behavior to internet access, and implement the quantitative analysis.

We proceed in two steps. First, we estimate the parameters that govern how fiber connectivity affects communication and information acquisition costs,namely, the communication-cost elasticity ($\gamma/\lambda$), the intercept of the information cost function ($\lambda\bar{\eta}$), and its slope with respect to connectivity ($\lambda\eta$). Because the outer-nest entropy weight $\lambda$ enters only as a common scale absorbed by fixed effects, only these composites are identified from the data. These parameters are central to the model's predictions and are identified using variation in fiber rollout across provinces and over time. Second, we calibrate the remaining structural parameters: the materials share of costs ($\alpha$), the consumption share of manufacturing goods ($\beta$), and the trade elasticity ($\zeta/\lambda$)—using standard values from production statistics and the trade literature.

## 5.1 Estimation of Fiber Connectivity Elasticities

The key parameters to be estimated, $\gamma/\lambda$, $\lambda\bar{\eta}$, and $\lambda\eta$, determine how improvements in fiber connectivity translate into reductions in communication and information acquisition costs. To identify

[7] See Appendix C.3 for details on the construction of the counterfactual equilibrium.

these parameters, we operationalize the supplier choice probabilities derived in Proposition 1 as a nested logit model. The nested structure reflects the two-stage nature of firms' sourcing decisions: first, the selection of a province from which to source inputs (the outer nest); and second, the selection of a specific supplier within that province (the inner nest).

### Estimation Framework

Making explicit the panel dimension of the data with subscript $t$, the nested logit probabilities take the following form:

$$\pi_{sd|o(s),t} = \frac{\exp\left(\ln \bar{\pi}_{sd|o(s)} - \zeta\left(\bar{\eta} + \eta \ln I_{o(s)d,t}\right)\ln c_{st}\right)}{\sum_{s' \in o(s)} \exp\left(\ln \bar{\pi}_{s'd|o(s)} - \zeta\left(\bar{\eta} + \eta \ln I_{o(s)d,t}\right)\ln c_{s't}\right)}, \tag{12}$$

$$\pi_{od,t} = \frac{\exp\left(\ln \bar{\pi}_{od} - \zeta \ln \tau_{od} + \gamma \ln I_{od,t} + IV_{od,t}\right)}{\sum_{o'} \exp\left(\ln \bar{\pi}_{o'd} - \zeta \ln \tau_{o'd} + \gamma \ln I_{o'd,t} + IV_{o'd,t}\right)}, \tag{13}$$

where the inclusive value is given by

$$IV_{od,t} = \frac{1}{\bar{\eta} + \eta \ln I_{od,t}} \ln\left(\sum_{s \in o} \exp\left(\ln \bar{\pi}_{sd|o} - \zeta\left(\bar{\eta} + \eta \ln I_{od,t}\right)\ln c_{st}\right)\right).$$

The inner nest probability in equation (12) characterizes the choice of a supplier conditional on having selected a province, while the outer nest probability in equation (13) governs the choice among provinces.

### Identification and Estimation Strategy

Estimation of this nested logit model requires controlling for several unobservable quantities: suppliers' marginal costs $c_{st}$, which vary across firms and over time, and the unconditional choice probabilities $\bar{\pi}_{sd|o}$ and $\bar{\pi}_{od}$, which capture firm- and market-specific demand shifters. From a fixed-effects perspective, these introduce a high-dimensional set of nuisance parameters.

To address the resulting computational challenges, we adopt a sequential estimation strategy. In the first stage, we estimate the inner nest, which governs firms' choice of suppliers within a province, while holding province-level characteristics constant. Because the inner nest exponent is linear in parameters, PPML directly yields reduced-form coefficients on $\ln c_{st}$ and on $\ln c_{st} \cdot \ln I_{od,t}$, whose ratio identifies $\eta/\bar{\eta}$. In the second stage, we estimate the outer nest, which governs the choice of province, using a *rescaled* inclusive value $\widetilde{IV}_{od,t}$ constructed entirely from first-stage estimates. Crucially, $\widetilde{IV}_{od,t}$ depends only on $\eta/\bar{\eta}$, not on $\bar{\eta}$ separately, because its prefactor approximates $\bar{\eta}\lambda_{od}$ to first order, so $\widetilde{IV}_{od,t} \approx \bar{\eta}\lambda_{od} \ln S_{od,t}$. The coefficient on this regressor in the outer nest then identifies $1/(\lambda\bar{\eta})$, where $\lambda$ is the outer-nest entropy weight. Because $\lambda$ enters the outer nest as a common scale factor absorbed by fixed effects, only the ratio $1/(\lambda\bar{\eta})$ is identified from the data; the level of $\lambda$ is a normalization. This two-step procedure follows standard approaches to nested logit estimation in settings with large choice sets and ensures consistency in the presence of unobserved heterogeneity.

To address potential endogeneity of fiber connectivity arising, for example, from correlation between fiber rollout and unobserved supply- or demand-side shocks, we employ a control function approach. Specifically, we include in both stages a control constructed from the residual of a first-stage regression that uses the instrumental variable described in Section 3. This strategy mitigates bias and allows us to recover consistent estimates of the target elasticities. Further details on the estimation procedure, including the construction of proxy variables (Section D.1), the PPML implementation (Section D.2), and the control function approach (Section D.3), are provided in Appendix D.

### 5.2 Calibration of Remaining Parameters

We now turn to the calibration of the remaining structural parameters: the trade elasticity $\zeta/\lambda$, the materials share $\alpha$, and the manufacturing consumption share $\beta$.

**Trade Elasticity.** Equation (9) establishes that trade costs enter the outer nest as $\tau_{od}^{-\zeta/\lambda}$, so the trade elasticity is $\zeta/\lambda = -\frac{\partial \ln(\pi_{od}/\pi_{dd})}{\partial \ln \tau_{od}}$. Following the quantitative trade literature, we set $\zeta/\lambda = 6.1$, consistent with the estimate reported by Adao et al. (2017). This value lies near the median of empirically estimated trade elasticities surveyed in Head and Mayer (2014).

**Production and Consumption Shares.** The materials share in manufacturing production, $\alpha$, is set to 0.7, corresponding to the average ratio of intermediate input expenditures to total costs across firms in our dataset. The consumption share of manufacturing goods in final demand, $\beta$, is set to 0.18 based on national expenditure data from the Turkish Statistical Institute (TUIK), ensuring consistency between the model's sectoral structure and the observed composition of household spending.

Table 4 summarizes all estimated and calibrated parameter values. Goodness-of-fit tests confirming the model's ability to replicate observed sourcing patterns are reported in Section D.6.

## 6 Welfare Effects of High-Speed Internet Infrastructure

With the model parameters estimated and calibrated in Section 5, we now turn to the quantitative analysis. This section solves for counterfactual equilibria under alternative fiber connectivity scenarios and evaluates the resulting welfare effects across provinces. To our knowledge, this exercise provides the first general equilibrium quantification of the welfare gains from ICT infrastructure improvements that arise through reductions in both communication and information acquisition costs.

We proceed in two steps. First, we construct the shocks to communication and information acquisition costs implied by the observed changes in fiber connectivity between 2012 and 2019.

Table 4: **Parameter Values**

| Parameter | | Value | Source |
|---|---|---|---|
| **Information Acquisition Costs**: | | | |
| Intercept | $\lambda\bar{\eta}$ | 0.666 | Estimated |
| | | (0.002) | |
| Elasticity w.r.t Fiber Connectivity | $\lambda\eta$ | 0.023 | Estimated |
| | | (0.008) | |
| **Communication Costs**: | | | |
| Elasticity w.r.t Fiber Connectivity | $\gamma/\lambda$ | 0.462 | Estimated |
| | | (0.024) | |
| Trade Elasticity | $\zeta/\lambda$ | 6.116 | Adao et al. (2017) |
| | | (0.948) | |
| **Materials Share of Costs** | $\alpha$ | 0.7 | Data |
| **Manufacturing Share of Consumption** | $\beta$ | 0.18 | TUIK |

**Note**: Standard errors are reported in parentheses for estimated parameters.

Second, we solve the equilibrium system using the hat algebra methodology to compute welfare changes, decomposing the total gains into the separate contributions of each channel.

### 6.1 Construction of Counterfactual Shocks

With the structural parameters estimated and calibrated, we now construct the shocks to communication and information acquisition costs that drive the counterfactual analysis. These shocks are based on the observed province-pair improvements in fiber connectivity realized between 2012 and 2019.

For each province pair $(o, d)$, we compute the change in fiber connectivity as the average change between the 2019 and 2012 levels:

$$\widehat{I}_{od} = \left(\frac{I_{od,2019}}{I_{od,2012}}\right)^{1/7}.$$

Using the estimated parameters from Section 5, we translate this connectivity improvement into shocks to the two friction channels.

First, the communication cost channel responds to connectivity through the scale parameter of the match-specific productivity distribution. Since $\phi_{od} = I_{od}^{\gamma}$ enters the outer nest as $\phi_{od}^{1/\lambda}$, the effective shock is:

$$\widehat{\phi}_{od}^{1/\lambda} = \widehat{I}_{od}^{\gamma/\lambda}.$$

A higher $\gamma/\lambda$ implies that communication frictions fall more steeply as connectivity improves, making suppliers in well-connected regions relatively more attractive.

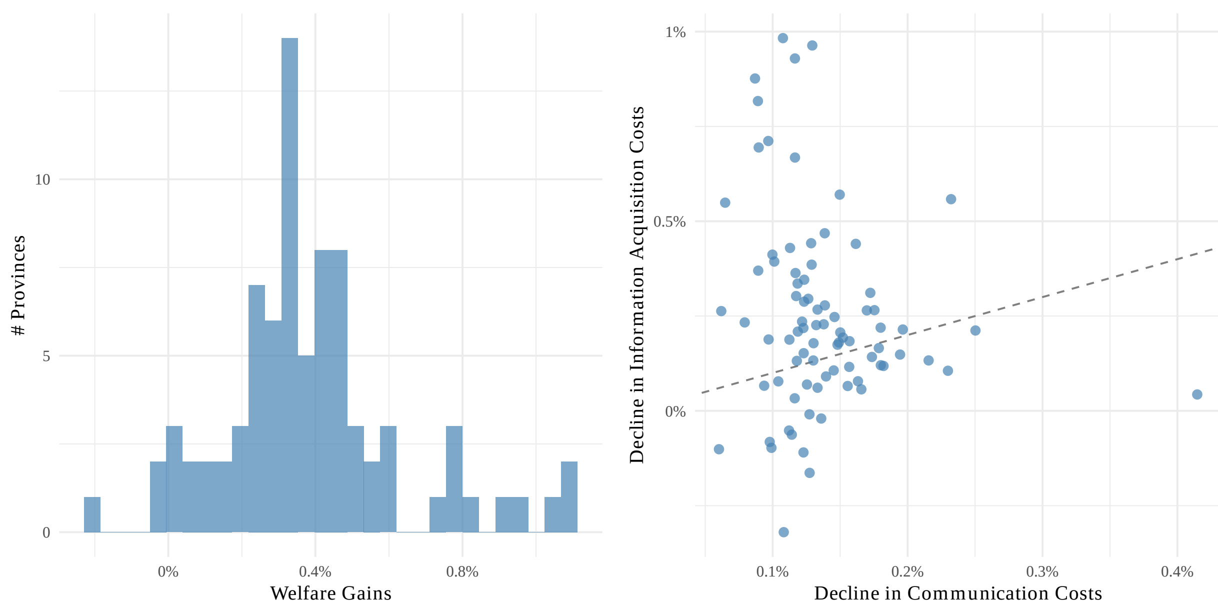
Second, the information acquisition cost channel responds through the cost-sensitivity ratio governing attention allocation. As connectivity improves, the ratio $\lambda_{od}/\lambda = 1/(\lambda\bar{\eta} + \lambda\eta \ln I_{od})$ falls, sharpening the buyer's responsiveness to cost differences among suppliers. This generates a shock $\widehat{\lambda}_{od}$ that varies with connectivity and the estimated parameter $\lambda\eta$.

Together, these two shocks provide the necessary exogenous variation to solve the counterfactual equilibrium system using the hat algebra methodology of Section 4. This approach allows us to isolate the effects of fiber rollout while holding all other economic fundamentals fixed at their 2012 levels.

### 6.2 Quantitative Results: Decomposing Welfare Gains

We take the Turkish economy in 2012 as the reference equilibrium and expose it to the province-pair improvements in fiber connectivity observed between 2012 and 2019. The counterfactual asks: how much would real income in each province have changed had the 2012 economy experienced these connectivity improvements, holding all other fundamentals fixed?

Figure 4: **Distribution of Provincial Welfare Gains and Channel Decomposition**




**Note:**The left panel shows the cross-provincial distribution of annualised welfare gains attributable to the 2012–2019 expansion of fibre infrastructure. The right panel plots each province's gain from the communication cost channel against its gain from the information acquisition cost channel, illustrating the relative contribution of each margin. Both panels report median gains across 200 bootstrap draws of the structural parameters.

To disentangle the channels through which improved connectivity affects welfare, we compute counterfactual equilibria that separately isolate reductions in communication costs and reductions in information acquisition costs. When both frictions decline jointly, we find an annualized median welfare gain of approximately 0.36 percent, with gains ranging from roughly 0.26 percent at the lower quartile to 0.48 percent at the upper quartile. Isolating each channel individually, we find a median gain of 0.13 percent from the decline in communication costs alone and 0.21 percent from the decline in information acquisition costs alone. These two contributions sum to 0.34 percentage points, slightly below the joint effect of 0.36 percent, indicating modest complementarity

Table 5: **Annualized Welfare Gains: Summary Statistics by Channel**

| | Welfare Changes | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Decline in: | | | |
| Communication Costs | ✓ | ✓ | |
| Information Acquisition Costs | ✓ | | ✓ |
| Mean | 0.4% | 0.14% | 0.26% |
| | [0.33%, 0.48%] | [0.11%, 0.18%] | [0.22%, 0.3%] |
| | (0.3%, 0.5%) | (0.1%, 0.2%) | (0.2%, 0.31%) |
| Lower Quartile | 0.26% | 0.12% | 0.11% |
| | [0.21%, 0.32%] | [0.09%, 0.15%] | [0.09%, 0.13%] |
| | (0.19%, 0.34%) | (0.08%, 0.17%) | (0.08%, 0.14%) |
| Median | 0.36% | 0.13% | 0.21% |
| | [0.29%, 0.43%] | [0.1%, 0.17%] | [0.18%, 0.26%] |
| | (0.27%, 0.46%) | (0.09%, 0.18%) | (0.16%, 0.27%) |
| Upper Quartile | 0.48% | 0.16% | 0.36% |
| | [0.4%, 0.59%] | [0.12%, 0.2%] | [0.3%, 0.43%] |
| | (0.37%, 0.61%) | (0.11%, 0.22%) | (0.27%, 0.44%) |

**Note:** Each entry reports the median annualised welfare gain across provinces, together with summary statistics of the cross-provincial distribution. Confidence intervals are obtained by bootstrapping the quantification exercise 200 times, drawing all structural parameters from a normal distribution centred at their point estimates with standard deviations equal to their standard errors. Numbers in round brackets are 95% confidence intervals; numbers in square brackets are 90% confidence intervals.

between the two channels.[8] Table 5 and Figure 4 report the full distribution of gains and their channel decomposition.[9]

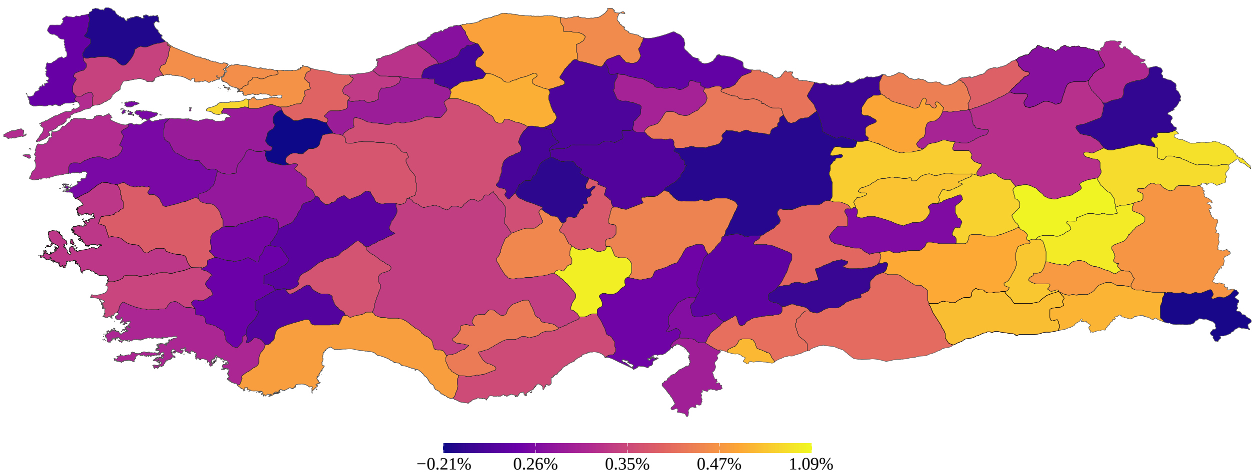
The annualized median welfare gain of approximately 0.36 percent corresponds to a cumulative gain of approximately 2.5 percent over the 2012–2019 period. For comparison, Coşar et al. (2022) estimate welfare gains from highway upgrading in Türkiye of around 2-3 percent, suggesting that digital infrastructure generates welfare returns of a comparable order of magnitude to major physical infrastructure investments.

Figure 5 maps the provincial distribution of these welfare gains. Panel (a) reveals that gains are broadly positive across the country but highly heterogeneous, reflecting the uneven geographic rollout of fiber infrastructure over the period. Provinces in the eastern and northeastern regions tend to record gains at the upper end of the distribution, reflecting larger reductions in bilateral information acquisition costs along their trade links. Notably, the rank correlation between a province's own fibre expansion and its welfare gain is only modest (0.23), underscoring that gains depend on the improvement in bilateral connectivity with trading partners rather than on a province's own infrastructure rollout alone. Panels (b) and (c) show that the two channels exhibit

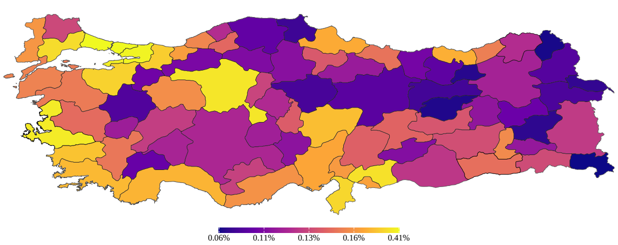
[8] Relative to the class of trade models studied by Arkolakis et al. (2012), our framework highlights an additional source of welfare gains: reductions in information acquisition costs. Because such frictions are absent in those models, the associated gains are necessarily omitted.

[9] To obtain confidence intervals, we bootstrap the quantification exercise 200 times. In each bootstrap, we draw the parameters listed in Table 4 from a normal distribution with a mean equal to their point estimate and a standard deviation equal to the standard error of the estimate.

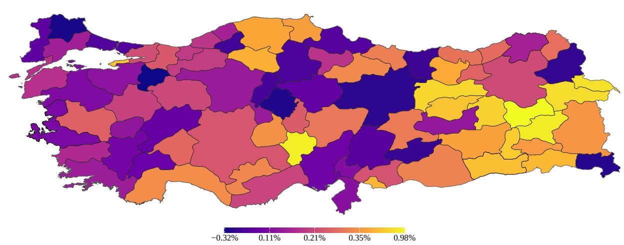
Figure 5: **Welfare Gains from Fibre Infrastructure Expansion across Turkish Provinces**




(a) Total Gain




(b) Communication Cost Channel




(c) Information Acquisition Cost Channel

**Note:** Each map plots the annualised welfare gain for each of the provinces attributable to the 2012–2019 expansion of fibre infrastructure. Panel (a) shows total gains. Panel (b) isolates the contribution of lower communication costs; panel (c) isolates the contribution of lower information acquisition costs. All gains are medians across 200 bootstrap draws of the structural parameters.

distinct spatial footprints. Gains from the decline in communication costs (panel b) are strongest in the large, economically central provinces — Istanbul, Kocaeli, İzmir, Ankara — where the volume of inter-provincial trade is greatest and reductions in communication costs translate into the largest trade share reallocations. By contrast, gains from the decline in information acquisition costs (panel c) exhibit substantially more heterogeneity across space (coefficient of variation of 1.03 versus 0.35 for the communication cost channel), with ten provinces recording negative gains along this margin, reflecting cases where connectivity improvements elsewhere made competing suppliers more attractive without a commensurate improvement in the province's own outward links.

Taken together, the results highlight that improved internet access raises welfare through two complementary mechanisms: lowering the costs of interacting and coordinating with suppliers (the communication channel) and enabling firms to better learn about potential suppliers (the information acquisition channel). Both mechanisms reduce effective input costs, but in different ways, and the largest welfare gains arise when provinces benefit from improvements along both margins. The geographic distribution of gains reflects underlying heterogeneity in initial frictions: gains from lower communication costs accrue primarily to the most connected, economically central provinces, whose larger trade volumes amplify the welfare effect of each bilateral cost reduction. By contrast, provinces that were initially more remote or informationally disadvantaged benefit disproportionately from reductions in information acquisition costs.[10]

Figure 6 provides a non-parametric validation of the model's cross-sectional predictions. Provinces in the upper quartile of predicted gains exhibit systematically higher median GDP growth than those in the bottom quartile. These patterns are consistent with a positive relationship between the extent to which a province stands to benefit from improved fibre connectivity and the GDP growth it subsequently realised.

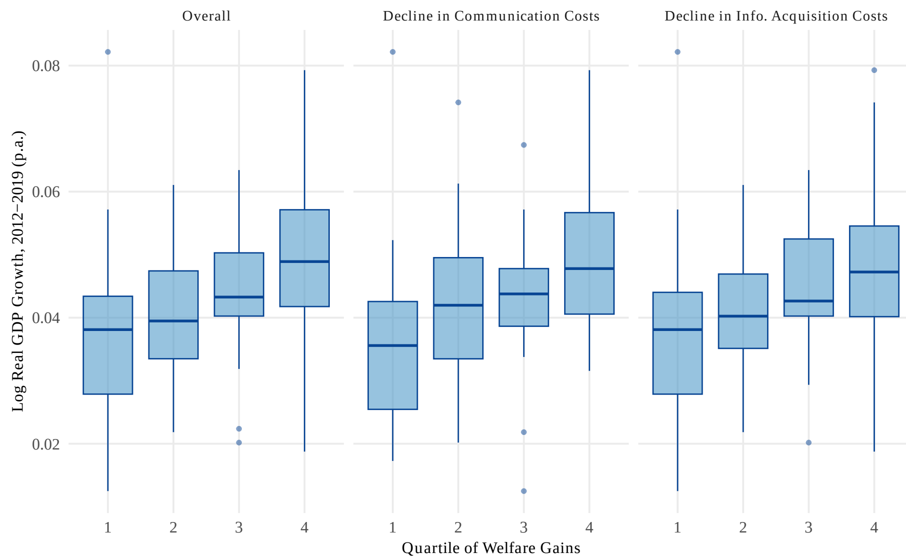
## 7 Conclusion

This paper provides the first causal evidence that high-speed internet infrastructure reshapes production networks and quantifies the resulting welfare gains in a general equilibrium framework. By analyzing the rollout of fiber-optic internet across Turkish provinces from 2012 to 2019 and leveraging rich microdata on firm-to-firm transactions, we document how improved connectivity transforms the spatial organization of supply chains.

Our empirical findings reveal two distinct patterns. First, firms reallocate their purchases toward suppliers located in provinces with better internet connectivity. Second, conditional on sourcing from a given province, firms diversify their supplier base, engaging with more suppliers and distributing purchases more evenly across them. These patterns are robust across a wide

[10] Across provinces, the welfare gains from lower communication costs are positively rank-correlated with initial fiber intensity (Spearman's $\rho = 0.78$) and negatively rank-correlated with travel time to Istanbul (Spearman's $\rho = -0.38$). By contrast, the welfare gains from lower information acquisition costs exhibit the opposite pattern: their rank correlation with initial fiber intensity is -0.32, while their rank correlation with travel time to Istanbul is 0.48.

Figure 6: **Predicted Welfare Gains and Realised GDP Growth**




**Note:** Each box plots actual annualised real GDP per capita growth over 2012–2019 for provinces grouped into quartiles of model-predicted welfare gains. Quartile assignments are computed separately for total gains, the communication cost channel, and the information acquisition cost channel. The horizontal line denotes the median; box edges span the interquartile range; whiskers extend to 1.5 $\times$ IQR; points beyond the whiskers are plotted individually.

range of specifications, including event-study analysis and placebo tests, and are identified using a novel instrument based on the pre-existing BOTAS pipeline network.

To rationalize these findings, we develop a general equilibrium model that embeds rational inattention into the endogenous formation of production networks. The model formalizes two channels through which internet connectivity affects sourcing decisions. Lower communication costs, arising from more reliable, high-bandwidth connections, increase the expected quality of match-specific productivity draws, making suppliers in well-connected provinces relatively more attractive and generating the reallocation pattern. Lower information acquisition costs enable firms to spread attention across a broader pool of potential suppliers rather than concentrating on a few known partners, generating the diversification pattern. The model yields testable predictions that align closely with the data.

Counterfactual simulations reveal that the expansion of fiber infrastructure raised real income by 2.5 percent in the median Turkish province. Reductions in information acquisition costs account for approximately 1.5 percentage points, while reductions in communication costs contribute about 0.9 percentage points. The magnitude of these gains is comparable to estimates on transportation infrastructure improvements.

Our findings carry direct policy implications. As governments and development banks invest heavily in broadband infrastructure, our results suggest that the returns extend beyond direct productivity gains to include the reorganization of supply chains, enabling firms to access better inputs, reduce sourcing costs, and build more resilient supplier networks. High-speed internet thus acts as a form of digital infrastructure that, like roads and ports, shapes the geography of production and promotes economic integration across space.

## References


ACEMOGLU, D., V. M. CARVALHO, A. OZDAGLAR, AND A. TAHBAZ-SALEHI (2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80, 1977–2016, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA9623.

ADAO, R., A. COSTINOT, AND D. DONALDSON (2017): "Nonparametric Counterfactual Predictions in Neoclassical Models of International Trade," *American Economic Review*, 107, 633–689.

AKER, J. C. (2010): "Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger," *American Economic Journal: Applied Economics*, 2, 46–59.

AKERMAN, A., I. GAARDER, AND M. MOGSTAD (2015): "The Skill Complementarity of Broadband Internet," *The Quarterly Journal of Economics*, 130, 1781–1824.

ALLEN, T. (2014): "Information Frictions in Trade," *Econometrica*, 82, 2041–2083, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10984.

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

ARKOLAKIS, C., A. COSTINOT, AND A. RODRÍGUEZ-CLARE (2012): "New Trade Models, Same Old Gains?" *American Economic Review*, 102, 94–130.

ARKOLAKIS, C., F. HUNEEUS, AND Y. MIYAUCHI (2025): "Production Network Formation, Trade, and Welfare," .

ATKIN, D. AND A. K. KHANDELWAL (2020): "How Distortions Alter the Impacts of International Trade in Developing Countries," *Annual Review of Economics*, 12, 213–238.

BAQAEE, D. R. AND E. FARHI (2024): "Networks, Barriers, and Trade," *Econometrica*, 92, 505–541, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA17513.

CALLAWAY, B., A. GOODMAN-BACON, AND P. H. C. SANT'ANNA (2024): "Event Studies with a Continuous Treatment," *AEA Papers and Proceedings*, 114, 601–605.

CALLAWAY, B. AND P. H. C. SANT'ANNA (2021): "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225, 200–230.

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

COŞAR, A. K., B. DEMIR, D. GHOSE, AND N. YOUNG (2022): "Road Capacity, Domestic Trade and Regional Outcomes," *Journal of Economic Geography*, 22, 901–929.

DANZER, A. M., C. FEUERBAUM, AND F. GAESSLER (2024): "Labor Supply and Automation Innovation: Evidence from an Allocation Policy," *Journal of Public Economics*, 235, 105136.

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. Y. XU, AND K. K. YANG (2024): "O-Ring Production Networks," *Journal of Political Economy*, 132, 200–247.

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

DINGEL, J. I. AND F. TINTELNOT (2020): "Spatial Economics for Granular Settings," .

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

DONALDSON, D. AND R. HORNBECK (2016): "Railroads and American Economic Growth: A "Market Access" Approach *," *The Quarterly Journal of Economics*, 131, 799–858.

EATON, J. AND S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

EATON, J., S. S. KORTUM, AND F. KRAMARZ (2022): "Firm-to-Firm Trade: Imports, Exports, and the Labor Market," .

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, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/iere.12469.

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

GOLDSMITH-PINKHAM, P., I. SORKIN, AND H. SWIFT (2020): "Bartik Instruments: What, When, Why, and How," *American Economic Review*, 110, 2586–2624.

HEAD, K. AND T. MAYER (2014): "Chapter 3 - 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, 1032–1079.

HJORT, J. AND L. TIAN (2025): "The Economic Impact of Internet Connectivity in Developing Countries," *Annual Review of Economics*, 17, 99–124.

JENSEN, R. (2007): "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector*," *The Quarterly Journal of Economics*, 122, 879–924.

JIANG, X. (2023): "Information and Communication Technology and Firm Geographic Expansion," *CESifo Working Paper Series*, number: 10452.

JUHÁSZ, R. AND C. STEINWENDER (2018): "Spinning the Web: The Impact of ICT on Trade in Intermediates and Technology Diffusion," .

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.

MATĚJKA, F. AND A. MCKAY (2015): "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105, 272–298.

MAZET-SONILHAC, C. (2021): "Search Frictions in Credit Markets," .

OBERFIELD, E. (2018): "A Theory of Input–Output Architecture," *Econometrica*, 86, 559–589, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10731.

PANIGRAHI, P. (2021): "Endogenous Spatial Production Networks: Quantitative Implications for Trade and Productivity," *CESifo Working Paper Series*, number: 9466.

RAMONDO, N., A. RODRÍGUEZ-CLARE, AND M. SABORÍO-RODRÍGUEZ (2016): "Trade, Domestic Frictions, and Scale Effects," *American Economic Review*, 106, 3159–3184.

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

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

SOTELO, S. (2020): "Domestic Trade Frictions and Agriculture," *Journal of Political Economy*, 128, 2690–2738.

TOMBE, T. AND X. ZHU (2019): "Trade, Migration, and Productivity: A Quantitative Analysis of China," *American Economic Review*, 109, 1843–1872.

# Online Appendix

## A Appendix: Background & Data

This appendix provides supplementary figures and tables for the background and data discussion in Section 2. We present additional visualizations of the growth of fiber internet subscriptions, the instrumental variable construction, the geographic variation in bilateral fiber connectivity, firm-level adoption patterns, and validation exercises linking fiber infrastructure to internet quality, adoption, and firm behavior.

### A.1 Growth of Fiber Internet

Figure A1 documents the rapid expansion of fiber internet in Türkiye over the sample period. The number of fiber subscribers increased five-fold between 2012 and 2019, and fiber's share of total fixed broadband connections rose substantially. Figure A2 shows changes in fiber length across provinces.

### A.2 Instrumental Variable: BOTAS Pipeline Network

As discussed in Section 2, our instrumental variable exploits the historical geography of Türkiye's natural gas infrastructure. Figure A3 displays the BOTAS oil and gas pipeline network, which also carries optical fiber cables. Figure A4 shows the spatial variation in population-weighted distance to this network across provinces, as defined in equation (1).

### A.3 Bilateral Fiber Connectivity

Figure A5 illustrates the change in bilateral fiber connectivity between each pair of Turkish provinces over the sample period. This bilateral measure, defined in Section 3 as the minimum of fiber intensity between provinces in a pair, captures the "weakest-link" nature of synchronous communication.

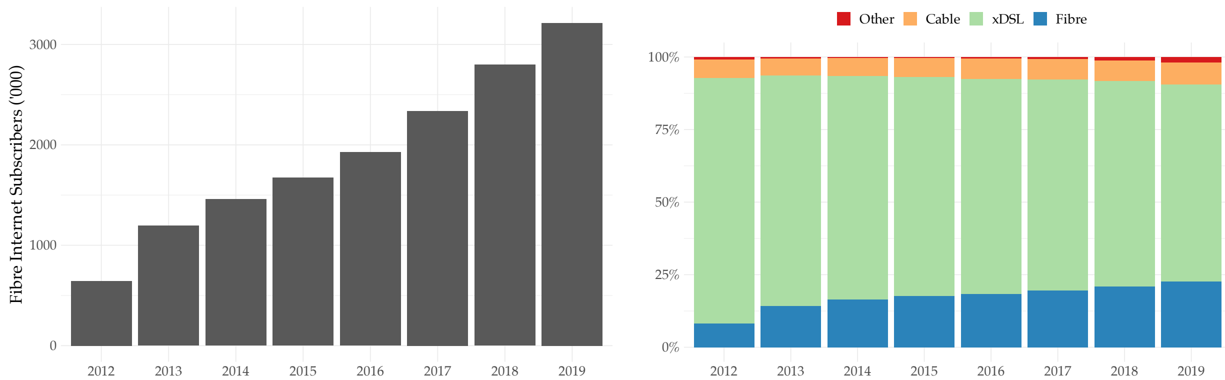
### A.4 Firm Adoption of High-Speed Internet

Figure A6 shows the evolution of firms' adoption of high-speed internet across broad sectors over the sample period. This figure complements the analysis in Section 2.6, which documents a strong positive relationship between provincial fiber rollout and firm-level adoption of connections exceeding 100 Mbps.

### A.5 Validation Tables

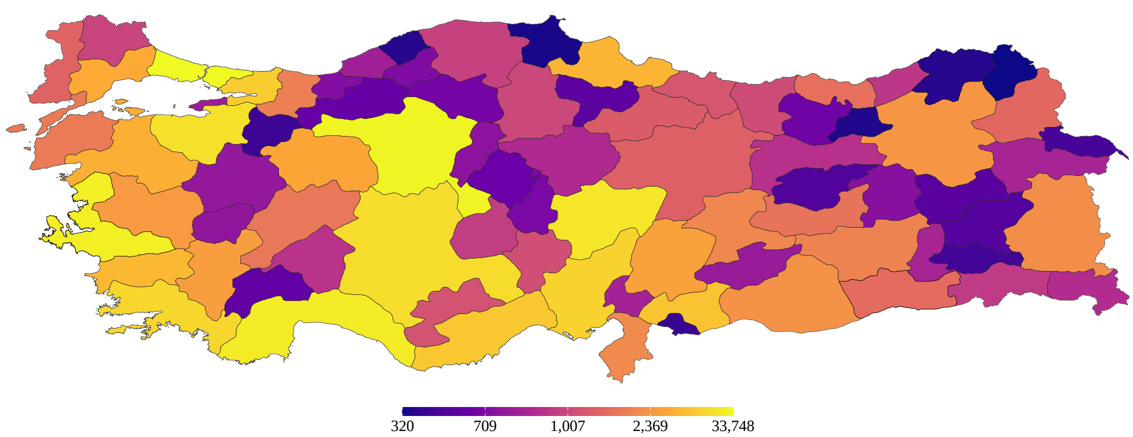
The following tables present additional results from the empirical validation exercises discussed in Section 2.6.

Figure A1: **Fiber Internet Subscribers**




**Note**: The left panel depicts the evolution of the number of fiber internet subscribers in Türkiye during the period 2012-2019. The right panel shows the breakdown of fixed broadband connections into fiber, 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 fiber internet also increased.

Figure A2: **Change in Optical Fiber Length**




**Note**: The map displays the change in total fibre cable length (km) by province between 2012 and 2019. Colors encode the percentile rank of the change across provinces, with lighter shades indicating larger increases. Legend labels show the actual kilometer values at the 0th, 25th, 50th, 75th, and 100th percentiles of the distribution.

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


**Note**: This map shows the gas pipeline network of BOTAS. Optical fiber cables were laid alongside these pipelines prior to the 2011 deregulation reform.

**Internet Quality.** Table A1 reports 2SLS estimates of the effect of fiber connectivity on internet speed and reliability, using Ookla's province-level data for 2016–2019.

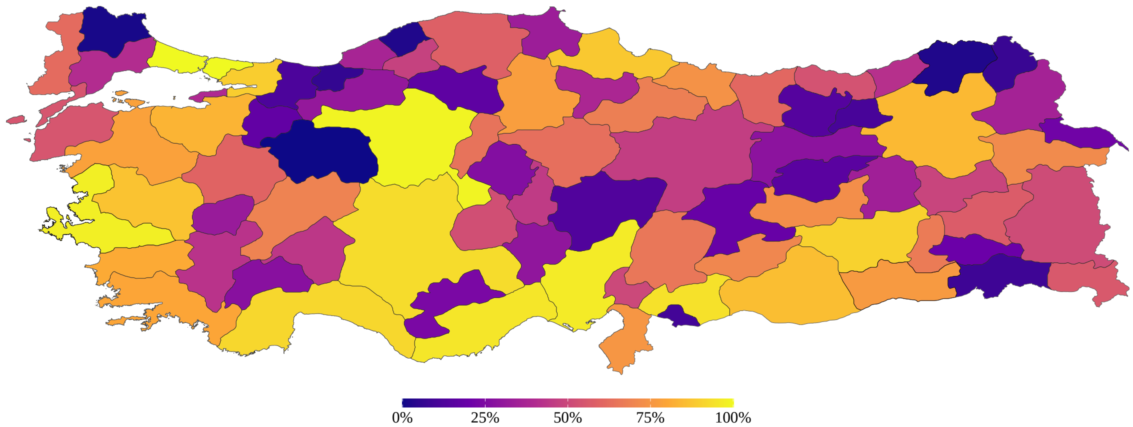
Table A1: **Quality of Internet Connection**

| Dependent Variable: | Download Speed | | Upload Speed | |
|---|---|---|---|---|
| | Mean (1) | Variance (2) | Mean (3) | Variance (4) |
| Fiber Connectivity | 2.396** (1.008) | -6.718** (2.801) | 4.793*** (1.816) | -8.860** (4.414) |
| **Fixed Effects & Controls:** | | | | |
| Province | ✓ | ✓ | ✓ | ✓ |
| Year | ✓ | ✓ | ✓ | ✓ |
| Initial Prov. Char. ×Year | ✓ | ✓ | ✓ | ✓ |
| Observations | 324 | 324 | 324 | 324 |

**Note**: See Section 2.6 for discussion. The dependent variables are average download speed, average upload speed, and within-province variance of speeds. All specifications include province and year fixed effects and controls for baseline provincial characteristics interacted with year dummies. Fiber Connectivity is instrumented with distance to BOTAS network. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered by province, in parentheses.

**Firm-Level Adoption.** Table A2 reports 2SLS estimates of the effect of provincial fiber intensity on firm-level adoption of high-speed internet (connections exceeding 100 Mbps), using the annual ICT Usage Survey.

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




**Note**: The map displays the population-weighted distance from each province to the nearest BOTAS natural gas pipeline network. Colors encode the percentile rank of the distance across provinces, with lighter shades indicating greater distance from the pipeline. This variable serves as an instrument for fibre internet infrastructure rollout. 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 (see equation (1)).

**Firm Behavior.** Table A3 reports 2SLS estimates linking fiber connectivity to changes in firm behavior, including average distance to suppliers, reliance on wholesalers, and purchases from ICT-related industries.

Table A2: **Firm Adoption of High-Speed Internet**

| Dependent Variable: | $\mathbb{I}\{$Firm has High-Speed Internet$\}$ | |
|---|---|---|
| | (1) OLS | (2) 2SLS |
| Fiber Connectivity | 0.410** (0.200) | 0.371* (0.216) |
| **Fixed Effects & Controls:** | | |
| Sector×Year | ✓ | ✓ |
| Province×Sector | ✓ | ✓ |
| Size class ×Year | ✓ | ✓ |
| Initial Prov. Chars.×Year | ✓ | ✓ |
| Observations | 1,218 | 1,218 |

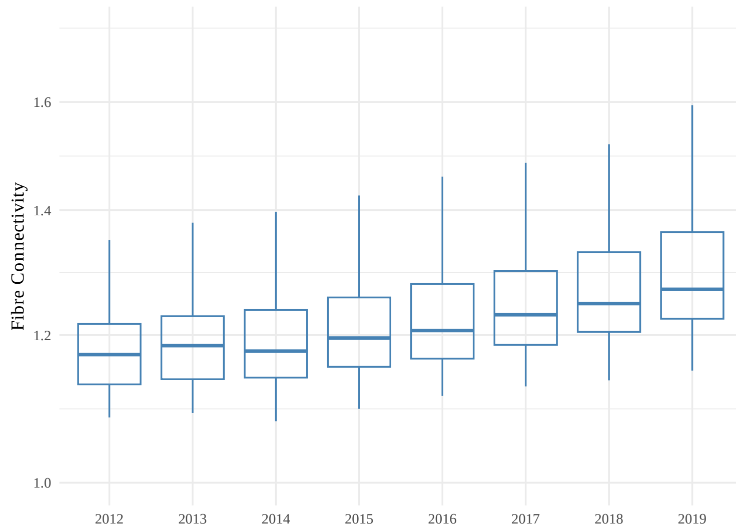
**Note**: See Section 2.6 for discussion. The dependent variable is an indicator for whether the firm reports a connection exceeding 100 Mbps. The specification controls for firm size and includes province and sector-year fixed effects (see equation (3)). * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered by province, in parentheses.

Table A3: **Impact of Fiber Connectivity on Firm Behavior**

| Dependent Variable: | Distance to Suppliers (Average) | Purchases from Wholesalers (Shares) | Purchases from ICT Firms (Log) |
|---|---|---|---|
| | (1) 2SLS | (2) 2SLS | (3) 2SLS |
| Fiber Connectivity | 0.0658** (0.0258) | -0.00842** (0.00348) | 0.143** (0.0548) |
| **Fixed Effects & Controls:** | | | |
| Firm | ✓ | ✓ | ✓ |
| Size class×Year | ✓ | ✓ | ✓ |
| Initial Prov. Chars. ×Year | ✓ | ✓ | ✓ |
| Observations | 541237 | 541237 | 541237 |

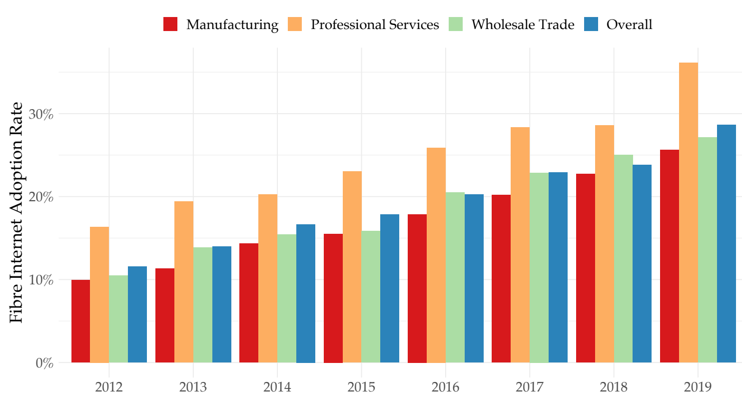
**Note**: See Section 2.6 for discussion. The dependent variables are: (i) average geographic distance to suppliers, (ii) share of input purchases from wholesalers, and (iii) total purchases from ICT-related industries. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered by province, in parentheses.

Figure A5: **Change in Bilateral Fiber Connectivity**




**Note:** This figure depicts a box and whiskers plot of the distribution of fiber connectivity across province pairs during the period 2012–2019. Fiber connectivity is measured as the minimum of fiber intensity between provinces in a pair.

Figure A6: **Firms' Adoption of High-Speed Internet in Türkiye**




**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 adopted high-speed internet. The fraction of firms with high-speed internet increased from zero in 2011 to about 30% by the end of 2019.

## B Appendix: Empirical Evidence

This appendix provides supplementary figures and tables for the empirical analysis presented in Section 3. Figure A5 in Appendix A illustrates the change in bilateral fiber connectivity between Turkish provinces over the sample period. Below we present the full set of robustness checks discussed in Section 3.2, followed by event-study estimates.

### B.1 Robustness Tables

The following tables present the robustness checks discussed in Section 3.2. All specifications follow the baseline equation (5) with modifications as noted.

**Alternative Measure of Connectivity.** Table A4 reports estimates using an alternative measure of bilateral fiber connectivity based on the similarity of digital infrastructure between trading partners, defined as $\ln \widetilde{I}_{odt} = -|\ln \widehat{I}_{ot} - \ln \widehat{I}_{dt}|$.

Table A4: **Alternative Measure of Bilateral Connectivity**

| Dependent Variable: | Cost Share (1) | No. Suppliers (2) | Cost Share HHI (3) | New Connections (4) |
|---|---|---|---|---|
| Fiber Connectivity, Alternative | 0.214*** (0.0235) | 0.131*** (0.0134) | -0.0423*** (0.00592) | 0.0244*** (0.00877) |
| **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 located in a given origin province. *New Connections* is the number of new suppliers relative to the year before. Fibre connectivity is measured as the negative absolute difference of log predicted fiber intensity at origin and destination provinces: $-|\ln \widehat{I}_{ot} - \ln \widehat{I}_{dt}|$, where $\widehat{I}_{o,t}$ denotes the instrument-predicted fiber intensity for province $o$ in year $t$. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

**Additional Controls.** Table A5 includes controls for the absolute difference in GDP per capita between origin and destination provinces and for 3G/4G mobile connectivity.

Table A5: **Additional Controls**

| Dependent Variable: | Cost Share | | No. Suppliers | | Cost Share HHI | | New Connections | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Fiber Connectivity | 0.383*** | 0.441*** | 0.227*** | 0.360*** | -0.0701*** | -0.0830*** | 0.0394** | 0.0393** |
| | (0.0473) | (0.0518) | (0.0264) | (0.0290) | (0.0116) | (0.0133) | (0.0177) | (0.0179) |
| Difference in GDP p.c. | 0.270** | | 0.211*** | | -0.0888*** | | 0.0579** | |
| | (0.0827) | | (0.0430) | | (0.0203) | | (0.0248) | |
| Mobile Connectivity | | -0.156 | | 0.0126 | | -0.0184 | | 0.109 |
| | | (0.248) | | (0.117) | | (0.0649) | | (0.0688) |
| **Fixed Effects:** | | | | | | | | |
| Buyer $\times$ Year | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Origin $\times$ Year | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Origin $\times$ 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 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. Fiber Connectivity is constructed using the instrument's predicted values, as defined in (4). * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

**Source of Variation in Connectivity.** Table A6 examines heterogeneity in treatment effects based on the quartile distribution of fiber intensity, exploiting the fact that our bilateral measure is dominated by the province with lower connectivity.

Table A6: **Source of Variation in Fibre Intensity**

| Dependent Variable: | Cost Share (1) | No. Suppliers (2) | Cost Share HHI (3) | New Connections (4) |
|---|---|---|---|---|
| **Panel A** | | | | |
| Origin Fiber Intensity × | | | | |
| Destination Fiber Intensity in: | | | | |
| $2^{nd}$ Quartile | -0.126 | -0.750 | 0.244 | -0.446 |
| | (0.305) | (0.587) | (0.273) | (0.375) |
| $3^{rd}$ Quartile | 0.132 | 0.0962 | -0.0435 | 0.0548 |
| | (0.149) | (0.0644) | (0.0349) | (0.0419) |
| $4^{th}$ Quartile | 0.433*** | 0.264*** | -0.0854*** | 0.0495*** |
| | (0.0468) | (0.0269) | (0.0119) | (0.0175) |
| **Panel B** | | | | |
| Destination Fiber Intensity × | | | | |
| Origin Fiber Intensity in: | | | | |
| $2^{nd}$ Quartile | -0.554 | -0.178 | 0.0994 | -0.0177 |
| | (0.300) | (0.175) | (0.0988) | (0.120) |
| $3^{rd}$ Quartile | 0.0859 | 0.0998 | -0.0199 | -0.0071 |
| | (0.146) | (0.0680) | (0.0356) | (0.0460) |
| $4^{th}$ Quartile | 0.431*** | 0.263*** | -0.0856*** | 0.0490*** |
| | (0.0475) | (0.0269) | (0.0118) | (0.0175) |
| **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 located in a given origin province. *New Connections* is the number of new suppliers relative to the year before. Fiber intensities are constructed using the instrument's predicted values, as defined in (4). * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

**Including Non-Manufacturing Suppliers.** Table A7 expands the sample to include all supplier firms, including those in services, trade, and other sectors.

**Excluding Multi-Region Firms.** Table A8 excludes buyers and suppliers that operate in multiple provinces to rule out internal reallocation within multi-establishment firms.

**Long Differences Between 2011 and 2019.** Table A9 presents long-difference estimates comparing cumulative changes in input sourcing to cumulative changes in predicted fiber connectivity.

Table A7: **Including Non-Manufacturing Suppliers**

| Dependent Variable: | Cost Share (1) | No. Suppliers (2) | Cost Share HHI (3) | New Connections (4) |
|---|---|---|---|---|
| Fiber Connectivity | 0.576*** | 0.321*** | -0.0973*** | 0.113*** |
| | (0.0518) | (0.0270) | (0.0104) | (0.0166) |
| **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 located in a given origin province. *New Connections* is the number of new suppliers relative to the year before. Fiber Connectivity is constructed using the instrument's predicted values, as defined in (4). * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

Table A8: **Excluding Multi-Region Firms**

| Dependent Variable: | Cost Share (1) | No. Suppliers (2) | Cost Share HHI (3) | New Connections (4) |
|---|---|---|---|---|
| Fiber Connectivity | 0.468*** | 0.277*** | -0.0955*** | 0.0548*** |
| | (0.0553) | (0.0324) | (0.0144) | (0.0196) |
| **Fixed Effects:** | | | | |
| Buyer×Year | ✓ | ✓ | ✓ | ✓ |
| Origin×Year | ✓ | ✓ | ✓ | ✓ |
| Origin×Destination | ✓ | ✓ | ✓ | ✓ |
| Observations | 1,338,389 | 1,338,389 | 1,338,389 | 1,338,389 |

**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 located in a given origin province. *New Connections* is the number of new suppliers relative to the year before. Fiber Connectivity is constructed using the instrument's predicted values, as defined in (4). * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

Table A9: **Long Changes Between 2011–2019**

| Dependent Variable: $\Delta^{2011-2019}$ | Cost Share (1) | No. Suppliers (2) | Cost Share HHI (3) |
|---|---|---|---|
| Fiber Connectivity$_{t=2019}$ | 0.110*** (0.0365) | 0.167*** (0.0394) | -0.0446** (0.0178) |
| **Fixed Effects:** | | | |
| Buyer | ✓ | ✓ | ✓ |
| Origin | ✓ | ✓ | ✓ |
| Observations | 65,762 | 65,762 | 65,762 |

**Note**: See Section 3.2 for discussion. Each observation pertains to a buyer firm and an origin province. All variables are in period changes of 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 located in a given origin province. *New Connections* is the number of new suppliers relative to the year before. Fiber Connectivity is constructed using the instrument's predicted values, as defined in (4). * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

**Placebo Test: Cable TV Rollout.** Table A10 conducts a placebo test replacing fiber connectivity with cable-TV subscription rates, which share similar physical infrastructure but primarily served residential entertainment rather than business-to-business communication in Türkiye during this period.

Table A10: **Placebo Test**

| Dependent Variable: | Cost Share (1) | No. Suppliers (2) | Cost Share HHI (3) | New Connections (4) |
|---|---|---|---|---|
| Cable TV Connectivity | -2.210 (1.754) | -0.375 (0.906) | -0.0624 (0.441) | -0.315 (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 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.

**Controlling for Travel Time.** Table A11 includes interactions between inter-province travel times and year dummies to control for time-varying reductions in physical trade costs.

Table A11: **Controlling for Travel Time**

| Dependent Variable: | Cost Share (1) | No. Suppliers (2) | Cost Share HHI (3) | New Connections (4) |
|---|---|---|---|---|
| Fiber Connectivity | 0.229*** | 0.131*** | -0.0457*** | 0.0375** |
| | (0.0448) | (0.0255) | (0.0126) | (0.0183) |
| **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 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. Fiber Connectivity is constructed using the instrument's predicted values, as defined in (4). * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

**Excluding Origin–Destination Fixed Effects.** Table A12 estimates the specification without origin–destination fixed effects, allowing the bilateral travel-time variable to be identified directly.

Table A12: **Excluding Origin–Destination Fixed Effects**

| Dependent Variable: | Cost Share (1) | No. Suppliers (2) | Cost Share HHI (3) | New Connections (4) |
|---|---|---|---|---|
| Fiber 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 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. Fiber Connectivity is constructed using the instrument's predicted values, as defined in (4). * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

**Bilateral Instrument.** Table A13 estimates a specification that directly instruments bilateral connectivity using an explicit bilateral function of the underlying distance measures. Specifically, we construct $\overline{\text{Distance}}_{od} = \max\left\{\overline{\text{Distance}}_o, \overline{\text{Distance}}_d\right\}$, and interact it with year dummies.

Table A13: **Using Bilateral Instrument**

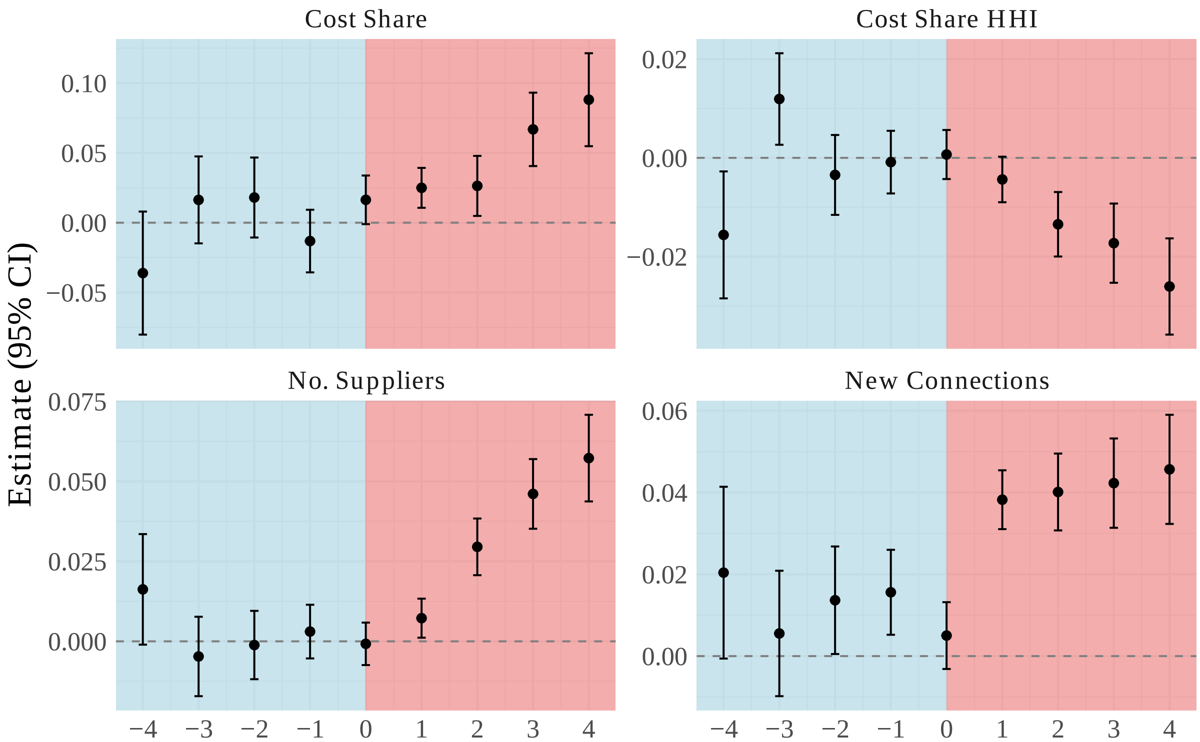
| Dependent Variable: | Cost Share<br>(1) | No. Suppliers<br>(2) | Cost Share HHI<br>(3) | New Connections<br>(4) |
|---|---|---|---|---|
| Fiber Connectivity | 0.498*** | 0.309*** | -0.102*** | 0.0629*** |
| | (0.0577) | (0.0310) | (0.0139) | (0.0197) |
| KP test stat. | 31.9 | 31.9 | 31.9 | 31.9 |
| **Fixed Effects:** | | | | |
| Buyer×Year | ✓ | ✓ | ✓ | ✓ |
| Origin×Year | ✓ | ✓ | ✓ | ✓ |
| Origin×Destination | ✓ | ✓ | ✓ | ✓ |
| Observations | 2,230,473 | 2,230,473 | 2,230,473 | 2,230,473 |

**Note**: 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 located in a given origin province. *New Connections* is the number of new suppliers relative to the year before. Fiber Connectivity is instrumented with $\overline{\text{Distance}}_{od} = \max\left\{\overline{\text{Distance}}_o, \overline{\text{Distance}}_d\right\}$, interacted with year dummies. * 10%, ** 5%, *** 1% significance levels. Standard errors, clustered at origin and destination level, are reported in parentheses.

## B.2 Event Study Estimates

Figure A7 presents event-study estimates for the four main outcome variables, providing visual evidence of the absence of pre-trends and the timing of treatment effects.

Figure A7: **Event Study Estimates**




**Note**: The figure shows event-study estimates based on the binary treatment variable defined in Section 3.2. The specification includes firm–source province fixed effects. Fiber Connectivity is constructed using the instrument's predicted values, as defined in (4).

# C Appendix: Model

This appendix provides formal proofs for the propositions stated in Section 4. For the reader's convenience, we restate each proposition before presenting its proof.

## C.1 Proof of Proposition 1 (Supplier Choice Probabilities)

**Proposition** (Restatement of Proposition 1)**.** *Conditional on the marginal cost of production for firm $s$ being $c_s$, the probability with which any firm located in $d$ selects supplier $s$ for any given task is*

$$\pi_{sd} = \pi_{sd|o(s)} \times \pi_{o(s)d},$$

$$\pi_{sd|o(s)} = \frac{\bar{\pi}_{sd|o(s)} c_s^{-\zeta/\lambda_{o(s)d}}}{\sum_{s' \in o(s)} \bar{\pi}_{s'd|o(s)} c_{s'}^{-\zeta/\lambda_{o(s)d}}}, \tag{14}$$

$$\pi_{od} = \frac{\bar{\pi}_{od} \tau_{od}^{-\zeta/\lambda} \phi_{od}^{1/\lambda} \left( \sum_{s \in o} \bar{\pi}_{sd|o} c_s^{-\zeta/\lambda_{od}} \right)^{\lambda_{od}/\lambda}}{\sum_{o'} \bar{\pi}_{o'd} \tau_{o'd}^{-\zeta/\lambda} \phi_{o'd}^{1/\lambda} \left( \sum_{s' \in o'} \bar{\pi}_{s'd|o'} c_{s'}^{-\zeta/\lambda_{o'd}} \right)^{\lambda_{o'd}/\lambda}}, \tag{15}$$

*where $\pi_{od}$ denotes the probability of choosing a supplier from province $o$ and $\pi_{sd|o(s)}$ the probability of choosing supplier $s$ conditional on having chosen to source from the province where it is located, $o(s)$.*

*Proof.* As discussed in the text, the buyer's sourcing problem is

$$\log p_d = \min_{\Pi_d \in \Delta} \left\{ \zeta \, \mathbb{E}_\Theta[\langle \Pi_d(\boldsymbol{v}_d), \boldsymbol{v}_d \rangle] + \psi(\Pi_d(\boldsymbol{v}_d), \Theta) \right\}, \tag{16}$$

where

$$\Delta = \left\{ \pi_{sd}(\boldsymbol{v}_d) \geq 0, \sum_s \pi_{sd}(\boldsymbol{v}_d) = 1 \right\},$$

and

$$\psi(\Pi_d(\boldsymbol{v}_d), \Theta) = \Omega(\bar{\Pi}_d) - \mathbb{E}_\Theta[\Omega(\Pi_d(\boldsymbol{v}_d))], \qquad \bar{\Pi}_d := \mathbb{E}_\Theta[\Pi_d(\boldsymbol{v}_d)].$$

The nested entropy functional is

$$\Omega(\Pi_d(\boldsymbol{v}_d)) = -\sum_o \left[ \lambda_{od} \sum_{s \in o} \pi_{sd}(\boldsymbol{v}_d) \ln \pi_{sd}(\boldsymbol{v}_d) + (\lambda - \lambda_{od}) \pi_{od}(\boldsymbol{v}_d) \ln \pi_{od}(\boldsymbol{v}_d) \right], \tag{17}$$

where

$$\pi_{od}(\boldsymbol{v}_d) := \sum_{s \in o} \pi_{sd}(\boldsymbol{v}_d).$$

**Notation and information structure.** The payoff-relevant state for supplier $s$ is its log effective cost $v_{sd} = \ln(c_s \tau_{o(s)d} / a_{sd})$, and $\boldsymbol{v}_d = (v_{sd})_s$ collects these across all suppliers. Here $c_s$ is the supplier's marginal cost (the object about which the buyer acquires information), $\tau_{od}$ is the bilateral communication cost (known to the buyer), and $a_{sd}$ is the match-specific productivity drawn i.i.d.

from a Fréchet distribution. The prior $\Theta$ denotes the buyer's joint distribution over $(\boldsymbol{c}, \boldsymbol{a}_d)$. Writing the policy as $\Pi_d(\boldsymbol{v}_d)$ allows the first-order conditions to be stated in full generality; the key simplification occurs in Step 5 below, where the i.i.d. match productivities are integrated out. Because $\mathbb{E}[\ln a_{sd}]$ is common across all suppliers $s$ within a given origin $o$, these terms cancel in within-origin comparisons, and the optimal posterior choice rule depends only on realized costs $\boldsymbol{c}$ and known bilateral trade costs $\boldsymbol{\tau}$. Thus, the expectation $\mathbb{E}_\Theta[\cdot]$ can equivalently be read as taken over cost uncertainty alone once the certainty-equivalent treatment of match productivity is applied.

**Step 1: Lagrangian and first-order conditions.** For $\psi(\Pi_d(\boldsymbol{v}_d), \Theta)$ convex and differentiable in $\Pi_d(\boldsymbol{v}_d)$, the Lagrangian is

$$\begin{aligned}\mathcal{L}(\Pi_d(\boldsymbol{v}_d), \mu_d(\boldsymbol{v}_d), \boldsymbol{\xi}_d(\boldsymbol{v}_d)) = \zeta \sum_s \int_{\boldsymbol{v}_d} \pi_{sd}(\boldsymbol{v}_d) v_{sd}\, \Theta(d\boldsymbol{v}_d) + \psi(\Pi_d(\boldsymbol{v}_d), \Theta) \\ + \int_{\boldsymbol{v}_d} \mu_d(\boldsymbol{v}_d) \left( \sum_s \pi_{sd}(\boldsymbol{v}_d) - 1 \right) \Theta(d\boldsymbol{v}_d) \\ - \int_{\boldsymbol{v}_d} \sum_s \xi_{sd}(\boldsymbol{v}_d) \pi_{sd}(\boldsymbol{v}_d)\, \Theta(d\boldsymbol{v}_d).\end{aligned}$$

The first-order and complementary-slackness conditions are

$$\frac{\partial \psi(\Pi_d(\boldsymbol{v}_d), \Theta)}{\partial \pi_{sd}(\boldsymbol{v}_d)} + \zeta v_{sd} + \mu_d(\boldsymbol{v}_d) - \xi_{sd}(\boldsymbol{v}_d) = 0, \tag{18}$$

$$\xi_{sd}(\boldsymbol{v}_d) \geq 0, \qquad \pi_{sd}(\boldsymbol{v}_d) \geq 0, \qquad \xi_{sd}(\boldsymbol{v}_d)\pi_{sd}(\boldsymbol{v}_d) = 0, \tag{19}$$

$$\sum_s \pi_{sd}(\boldsymbol{v}_d) = 1. \tag{20}$$

Hence, for any interior solution with $\pi_{sd}(\boldsymbol{v}_d) > 0$, complementary slackness implies $\xi_{sd}(\boldsymbol{v}_d) = 0$, and therefore

$$\frac{\partial \psi(\Pi_d(\boldsymbol{v}_d), \Theta)}{\partial \pi_{sd}(\boldsymbol{v}_d)} = -\mu_d(\boldsymbol{v}_d) - \zeta v_{sd}. \tag{21}$$

**Step 2: Gradient of the nested entropy functional.** We first compute the gradient of $\Omega(\cdot)$ for a generic $\Pi$. Using $\pi_{od}(\boldsymbol{v}_d) = \sum_{s \in o} \pi_{sd}(\boldsymbol{v}_d)$, we can write

$$\begin{aligned}\Omega(\Pi_d(\boldsymbol{v}_d)) &= -\sum_s \pi_{sd}(\boldsymbol{v}_d) \left[ \lambda_{o(s)d} \ln \pi_{sd}(\boldsymbol{v}_d) + (\lambda - \lambda_{o(s)d}) \ln \left( \sum_{s' \in o(s)} \pi_{s'd}(\boldsymbol{v}_d) \right) \right] \\ &= -\sum_o \sum_{s \in o} \pi_{sd}(\boldsymbol{v}_d) \left[ \lambda_{od} \ln \pi_{sd}(\boldsymbol{v}_d) + (\lambda - \lambda_{od}) \ln \left( \sum_{s' \in o} \pi_{s'd}(\boldsymbol{v}_d) \right) \right] \\ &= -\sum_o \left[ \lambda_{od} \sum_{s \in o} \pi_{sd}(\boldsymbol{v}_d) \ln \pi_{sd}(\boldsymbol{v}_d) + (\lambda - \lambda_{od}) \pi_{od}(\boldsymbol{v}_d) \ln \pi_{od}(\boldsymbol{v}_d) \right].\end{aligned}$$

Differentiating with respect to $\pi_{sd}(\boldsymbol{v}_d)$ gives

$$\frac{\partial\Omega(\Pi_d(\boldsymbol{v}_d))}{\partial\pi_{sd}(\boldsymbol{v}_d)} = -\lambda_{o(s)d}\big(\ln\pi_{sd}(\boldsymbol{v}_d)+1\big) - (\lambda-\lambda_{o(s)d})\big(\ln\pi_{od}(\boldsymbol{v}_d)+1\big). \tag{22}$$

Similarly,

$$\frac{\partial\Omega(\bar{\Pi}_d)}{\partial\bar{\pi}_{sd}} = -\lambda_{o(s)d}(\ln\bar{\pi}_{sd}+1) - (\lambda-\lambda_{o(s)d})(\ln\bar{\pi}_{od}+1). \tag{23}$$

**Step 3: Gradient of the information cost.** Using

$$\psi(\Pi_d(\boldsymbol{v}_d),\Theta) = \Omega(\bar{\Pi}_d) - \mathbb{E}_\Theta[\Omega(\Pi_d(\boldsymbol{v}_d))],$$

the derivative with respect to the state-contingent choice $\pi_{sd}(\boldsymbol{v}_d)$ is

$$\frac{\partial\psi(\Pi_d(\boldsymbol{v}_d),\Theta)}{\partial\pi_{sd}(\boldsymbol{v}_d)} = \frac{\partial\Omega(\bar{\Pi}_d)}{\partial\bar{\pi}_{sd}} - \frac{\partial\Omega(\Pi_d(\boldsymbol{v}_d))}{\partial\pi_{sd}(\boldsymbol{v}_d)}. \tag{24}$$

Substituting (22) and (23) into (24) yields

$$\begin{aligned}\frac{\partial\psi(\Pi_d(\boldsymbol{v}_d),\Theta)}{\partial\pi_{sd}(\boldsymbol{v}_d)} &= -\lambda_{o(s)d}(\ln\bar{\pi}_{sd}+1) - (\lambda-\lambda_{o(s)d})(\ln\bar{\pi}_{od}+1) \\ &\quad + \lambda_{o(s)d}(\ln\pi_{sd}(\boldsymbol{v}_d)+1) + (\lambda-\lambda_{o(s)d})(\ln\pi_{od}(\boldsymbol{v}_d)+1).\end{aligned} \tag{25}$$

**Step 4: State-contingent first-order condition.** Substituting (25) into (21), we obtain

$$\lambda_{o(s)d}\ln\pi_{sd}(\boldsymbol{v}_d) + (\lambda-\lambda_{o(s)d})\ln\pi_{od}(\boldsymbol{v}_d) = -\mu_d(\boldsymbol{v}_d) - \zeta v_{sd} + \lambda_{o(s)d}\ln\bar{\pi}_{sd} + (\lambda-\lambda_{o(s)d})\ln\bar{\pi}_{od}. \tag{26}$$

**Step 5: Supplier choice within an origin.** Fix an origin $o$, and take two suppliers $s, r \in o$. Subtracting the corresponding first-order conditions in (26) gives

$$\lambda_{od}\big(\ln\pi_{sd}(\boldsymbol{v}_d) - \ln\pi_{rd}(\boldsymbol{v}_d)\big) = -\zeta(v_{sd}-v_{rd}) + \lambda_{od}\big(\ln\bar{\pi}_{sd} - \ln\bar{\pi}_{rd}\big). \tag{27}$$

Exponentiating,

$$\frac{\pi_{sd}(\boldsymbol{v}_d)}{\pi_{rd}(\boldsymbol{v}_d)} = \frac{\bar{\pi}_{sd}e^{-\zeta v_{sd}/\lambda_{od}}}{\bar{\pi}_{rd}e^{-\zeta v_{rd}/\lambda_{od}}}. \tag{28}$$

Normalizing within origin $o$,

$$\frac{\pi_{sd}(\boldsymbol{v}_d)}{\pi_{od}(\boldsymbol{v}_d)} = \frac{\bar{\pi}_{sd}e^{-\zeta v_{sd}/\lambda_{od}}}{\sum_{s'\in o}\bar{\pi}_{s'd}e^{-\zeta v_{s'd}/\lambda_{od}}} = \frac{\bar{\pi}_{sd|o}e^{-\zeta v_{sd}/\lambda_{od}}}{\sum_{s'\in o}\bar{\pi}_{s'd|o}e^{-\zeta v_{s'd}/\lambda_{od}}}. \tag{29}$$

Now use $v_{sd} = \ln(c_s \tau_{od}/a_{sd})$. Since $\tau_{od}$ is common within origin $o$,

$$e^{-\zeta v_{sd}/\lambda_{od}} = \left(\frac{c_s \tau_{od}}{a_{sd}}\right)^{-\zeta/\lambda_{od}} \propto c_s^{-\zeta/\lambda_{od}} a_{sd}^{\zeta/\lambda_{od}}.$$

Under the certainty-equivalent treatment of match productivity,

$$\zeta E[\ln a_{sd}] = \ln \phi_{od} + \gamma_E,$$

which is common across all suppliers $s \in o$ and therefore cancels in the inner nest. Hence the conditional supplier choice probability reduces to

$$\pi_{sd|o} = \frac{\bar{\pi}_{sd|o} c_s^{-\zeta/\lambda_{od}}}{\sum_{r \in o} \bar{\pi}_{rd|o} c_r^{-\zeta/\lambda_{od}}}. \tag{30}$$

**Step 6: Origin choice probabilities.** Returning to (26), exponentiating gives

$$\pi_{sd}(\boldsymbol{v}_d)^{\lambda_{o(s)d}} \pi_{od}(\boldsymbol{v}_d)^{\lambda - \lambda_{o(s)d}} = e^{-\mu_d(\boldsymbol{v}_d)} e^{-\zeta v_{sd}} \bar{\pi}_{sd}^{\lambda_{o(s)d}} \bar{\pi}_{od}^{\lambda - \lambda_{o(s)d}}. \tag{31}$$

Rearranging,

$$\pi_{sd}(\boldsymbol{v}_d) = e^{-\zeta v_{sd}/\lambda_{o(s)d}} \bar{\pi}_{sd} \left[ e^{-\mu_d(\boldsymbol{v}_d)} \left( \frac{\bar{\pi}_{o(s)d}}{\pi_{o(s)d}(\boldsymbol{v}_d)} \right)^{\lambda - \lambda_{o(s)d}} \right]^{1/\lambda_{o(s)d}}. \tag{32}$$

Summing over all $s \in o$,

$$\begin{aligned} \pi_{od}(\boldsymbol{v}_d) &= \left( \sum_{s \in o} e^{-\zeta v_{sd}/\lambda_{od}} \bar{\pi}_{sd} \right) \left[ e^{-\mu_d(\boldsymbol{v}_d)} \left( \frac{\bar{\pi}_{od}}{\pi_{od}(\boldsymbol{v}_d)} \right)^{\lambda - \lambda_{od}} \right]^{1/\lambda_{od}} \\ &= \left( \sum_{s \in o} e^{-\zeta v_{sd}/\lambda_{od}} \bar{\pi}_{sd} \right)^{\lambda_{od}/\lambda} e^{-\mu_d(\boldsymbol{v}_d)/\lambda} \bar{\pi}_{od}^{(\lambda - \lambda_{od})/\lambda} \\ &= \bar{\pi}_{od} \left( \sum_{s \in o} \bar{\pi}_{sd|o} e^{-\zeta v_{sd}/\lambda_{od}} \right)^{\lambda_{od}/\lambda} e^{-\mu_d(\boldsymbol{v}_d)/\lambda}. \end{aligned} \tag{33}$$

Imposing $\sum_o \pi_{od}(\boldsymbol{v}_d) = 1$ gives

$$e^{\mu_d(\boldsymbol{v}_d)/\lambda} = \sum_o \bar{\pi}_{od} \left( \sum_{s \in o} \bar{\pi}_{sd|o} e^{-\zeta v_{sd}/\lambda_{od}} \right)^{\lambda_{od}/\lambda}. \tag{34}$$

Substituting back into (33),

$$\pi_{od}(\boldsymbol{v}_d) = \frac{\bar{\pi}_{od}\left(\sum_{s\in o}\bar{\pi}_{sd|o}e^{-\zeta v_{sd}/\lambda_{od}}\right)^{\lambda_{od}/\lambda}}{\sum_{o'}\bar{\pi}_{o'd}\left(\sum_{s\in o'}\bar{\pi}_{sd|o'}e^{-\zeta v_{sd}/\lambda_{o'd}}\right)^{\lambda_{o'd}/\lambda}}. \tag{35}$$

Finally, using $v_{sd} = \ln(c_s\tau_{od}/a_{sd})$ and the certainty-equivalent representation

$$\zeta E[\ln a_{sd}] = \ln\phi_{od} + \gamma_E,$$

we obtain

$$e^{-\zeta v_{sd}/\lambda_{od}} = c_s^{-\zeta/\lambda_{od}}\tau_{od}^{-\zeta/\lambda_{od}}\exp\left(\frac{\ln\phi_{od}+\gamma_E}{\lambda_{od}}\right),$$

so that

$$\left(\sum_{s\in o}\bar{\pi}_{sd|o}e^{-\zeta v_{sd}/\lambda_{od}}\right)^{\lambda_{od}/\lambda} = \tau_{od}^{-\zeta/\lambda}\phi_{od}^{1/\lambda}e^{\gamma_E/\lambda}\left(\sum_{s\in o}\bar{\pi}_{sd|o}c_s^{-\zeta/\lambda_{od}}\right)^{\lambda_{od}/\lambda}. \tag{36}$$

Because $e^{\gamma_E/\lambda}$ is common across all origins, it cancels in the normalization. Therefore

$$\pi_{od} = \frac{\bar{\pi}_{od}\tau_{od}^{-\zeta/\lambda}\phi_{od}^{1/\lambda}\left(\sum_{s\in o}\bar{\pi}_{sd|o}c_s^{-\zeta/\lambda_{od}}\right)^{\lambda_{od}/\lambda}}{\sum_{o'}\bar{\pi}_{o'd}\tau_{o'd}^{-\zeta/\lambda}\phi_{o'd}^{1/\lambda}\left(\sum_{s\in o'}\bar{\pi}_{sd|o'}c_s^{-\zeta/\lambda_{o'd}}\right)^{\lambda_{o'd}/\lambda}}. \tag{37}$$

Combining (30) and (37), the overall supplier choice probability is

$$\pi_{sd} = \pi_{sd|o(s)}\pi_{o(s)d},$$

which proves the proposition. □

### C.2 Proof of Proposition 2 (Indirect Utility)

**Proposition** (Restatement of Proposition 2). *The indirect utility is given by:*

$$V_d = \left(\frac{w_d}{P_d}\right)^{\beta} T_d^{1-\beta},$$

*where $P_d$ is given by:*

$$P_d = \exp\left(-\frac{\lambda}{\zeta}\mathbb{E}_{\Theta}\left[\ln\left(\sum_{o}\bar{\pi}_{od}\tau_{od}^{-\zeta/\lambda}\phi_{od}^{1/\lambda}\left(\sum_{s\in o}\bar{\pi}_{sd|o}c_s^{-\zeta/\lambda_{od}}\right)^{\lambda_{od}/\lambda}\right)\right]\right)$$

*Proof.* The proof proceeds in three steps.

**Step 1: Household expenditure minimization.** Households in province $d$ consume tradable manufacturing goods, denoted $C_d$, and a local non-traded composite, denoted $C_d^*$. Preferences are Cobb–Douglas:

$$u_d = \left(\frac{C_d}{\beta}\right)^{\beta}\left(\frac{C_d^*}{1-\beta}\right)^{1-\beta}.$$

The price of the local non-traded good is $p_d^* = \frac{w_d}{T_d}$ because the non-traded sector uses only local labor under constant returns to scale and prices at marginal cost.

Let $P_d$ denote the price index of the tradable manufacturing composite. Then the expenditure minimization problem is

$$e_d(u) = \min_{C_d, C_d^*}\left\{P_d C_d + \frac{w_d}{T_d}C_d^* : \left(\frac{C_d}{\beta}\right)^{\beta}\left(\frac{C_d^*}{1-\beta}\right)^{1-\beta} \geq u\right\}.$$

For normalized Cobb–Douglas preferences, the corresponding expenditure function is

$$e_d(u) = u\, P_d^{\beta}\left(\frac{w_d}{T_d}\right)^{1-\beta}.$$

Each household supplies one unit of labor inelastically and earns nominal income $w_d$. Therefore indirect utility is

$$V_d = \frac{w_d}{e_d(1)} = \frac{w_d}{P_d^{\beta}(w_d/T_d)^{1-\beta}} = \left(\frac{w_d}{P_d}\right)^{\beta} T_d^{1-\beta}.$$

Thus it remains to characterize $P_d$.

**Step 2: Task-level tradable price.** By assumption, the household's supplier choice problem for a manufacturing task is analogous to the firm's sourcing problem. Hence the unit price paid by households for one tradable task equals the optimized value of the same rational-inattention problem that governs firm input sourcing. Let

$$v_{sd} = \ln\left(\frac{c_s \tau_{od}}{a_{sd}}\right), \qquad o = o(s).$$

Under the form used in Proposition 1, the optimized value for one task can be written as

$$\zeta \ln p_d(\Theta) = \min_{\Pi_d \in \Delta}\left\{\zeta\, \mathbb{E}_{\Theta}\left[\sum_s \pi_{sd}(\boldsymbol{v}_d) v_{sd}\right] + \psi(\Pi_d(\boldsymbol{v}_d), \Theta)\right\}. \tag{38}$$

We write $\zeta \ln p_d$ on the left-hand side because the payoffs $v_{sd}$ and information costs $\psi$ are both measured in units of $\zeta \times$ (log price); dividing through by $\zeta$ at the end recovers the actual log-price. The first-order conditions are identical to those in Proposition 1. Hence the state-contingent origin choice probabilities satisfy

$$\pi_{od}(\boldsymbol{v}_d) = \frac{\bar{\pi}_{od}\left(\sum_{s\in o}\bar{\pi}_{sd|o}e^{-\zeta v_{sd}/\lambda_{od}}\right)^{\lambda_{od}/\lambda}}{\sum_{o'}\bar{\pi}_{o'd}\left(\sum_{s\in o'}\bar{\pi}_{sd|o'}e^{-\zeta v_{sd}/\lambda_{o'd}}\right)^{\lambda_{o'd}/\lambda}}. \tag{39}$$

Equivalently, the multiplier on the simplex constraint satisfies

$$e^{\mu_d(\boldsymbol{v}_d)/\lambda} = \sum_o \bar{\pi}_{od}\left(\sum_{s\in o}\bar{\pi}_{sd|o}e^{-\zeta v_{sd}/\lambda_{od}}\right)^{\lambda_{od}/\lambda},$$

so the minimized value of the problem is

$$\zeta \ln p_d(\Theta) = -\mu_d(\boldsymbol{v}_d) = -\lambda \ln\left[\sum_o \bar{\pi}_{od}\left(\sum_{s\in o}\bar{\pi}_{sd|o}e^{-\zeta v_{sd}/\lambda_{od}}\right)^{\lambda_{od}/\lambda}\right]. \tag{40}$$

**Step 3: Substitute the Fréchet structure and aggregate across tasks.** Using $v_{sd} = \ln\left(\frac{c_s\tau_{od}}{a_{sd}}\right)$, we have

$$e^{-\zeta v_{sd}/\lambda_{od}} = c_s^{-\zeta/\lambda_{od}}\tau_{od}^{-\zeta/\lambda_{od}}a_{sd}^{\zeta/\lambda_{od}}.$$

Because households do not acquire information about match-specific productivities, we use the same certainty-equivalent representation as in Proposition 1. Under

$$\Pr(a_{sd} \le a) = \exp(-\phi_{od}a^{-\zeta}),$$

we have

$$E[\ln a_{sd}] = \frac{\ln\phi_{od} + \gamma_E}{\zeta},$$

so

$$\zeta E[\ln a_{sd}] = \ln\phi_{od} + \gamma_E.$$

Therefore,

$$e^{-\zeta v_{sd}/\lambda_{od}} = c_s^{-\zeta/\lambda_{od}}\tau_{od}^{-\zeta/\lambda_{od}}\exp\left(\frac{\ln\phi_{od}+\gamma_E}{\lambda_{od}}\right),$$

and hence

$$\left(\sum_{s\in o}\bar{\pi}_{sd|o}e^{-\zeta v_{sd}/\lambda_{od}}\right)^{\lambda_{od}/\lambda} = \tau_{od}^{-\zeta/\lambda}\phi_{od}^{1/\lambda}e^{\gamma_E/\lambda}\left(\sum_{s\in o}\bar{\pi}_{sd|o}c_s^{-\zeta/\lambda_{od}}\right)^{\lambda_{od}/\lambda}. \tag{41}$$

Substituting (41) into (40) gives

$$\zeta \ln p_d(\Theta) = -\lambda \ln\left[e^{\gamma_E/\lambda}\sum_o \bar{\pi}_{od}\tau_{od}^{-\zeta/\lambda}\phi_{od}^{1/\lambda}\left(\sum_{s\in o}\bar{\pi}_{sd|o}c_s^{-\zeta/\lambda_{od}}\right)^{\lambda_{od}/\lambda}\right]$$

$$= -\gamma_E - \lambda \ln \left[ \sum_o \bar{\pi}_{od} \tau_{od}^{-\zeta/\lambda} \phi_{od}^{1/\lambda} \left( \sum_{s \in o} \bar{\pi}_{sd|o} c_s^{-\zeta/\lambda_{od}} \right)^{\lambda_{od}/\lambda} \right]. \tag{42}$$

The additive constant $-\gamma_E$ is common across all provinces and can be absorbed into the normalization of the tradable price index. Dividing both sides by $\zeta$ to recover the actual log-price,

$$\ln p_d(\Theta) = -\frac{\lambda}{\zeta} \ln \left[ \sum_o \bar{\pi}_{od} \tau_{od}^{-\zeta/\lambda} \phi_{od}^{1/\lambda} \left( \sum_{s \in o} \bar{\pi}_{sd|o} c_s^{-\zeta/\lambda_{od}} \right)^{\lambda_{od}/\lambda} \right]. \tag{43}$$

Finally, the tradable composite is Cobb–Douglas across tasks. Hence its exact price index is the geometric mean of task-level prices. Taking expectations over $\Theta$ therefore yields

$$\ln P_d = E_\Theta[\ln p_d(\Theta)] = -\frac{\lambda}{\zeta} E_\Theta \left[ \ln \left( \sum_o \bar{\pi}_{od} \tau_{od}^{-\zeta/\lambda} \phi_{od}^{1/\lambda} \left( \sum_{s \in o} \bar{\pi}_{sd|o} c_s^{-\zeta/\lambda_{od}} \right)^{\lambda_{od}/\lambda} \right) \right].$$

Exponentiating,

$$P_d = \exp \left( -\frac{\lambda}{\zeta} E_\Theta \left[ \ln \left( \sum_o \bar{\pi}_{od} \tau_{od}^{-\zeta/\lambda} \phi_{od}^{1/\lambda} \left( \sum_{s \in o} \bar{\pi}_{sd|o} c_s^{-\zeta/\lambda_{od}} \right)^{\lambda_{od}/\lambda} \right) \right] \right).$$

Combining this expression with Step 1 yields

$$V_d = \left( \frac{w_d}{P_d} \right)^{\beta} T_d^{1-\beta},$$

as claimed. □

### C.3 Solving for the Counterfactual Equilibrium

In the counterfactual analysis, we work with the large market aggregate representation of Proposition 2. Proposition 2 writes the household price index as an expectation over buyers' prior information. In a large economy, however, this expectation admits a deterministic aggregate counterpart when we assume that buyer-specific prior distortions average out in the cross section, so the aggregate prior converges to the uniform prior within each origin. Once Proposition 1 is written in terms of these aggregate prior objects, the term inside the expectation in Proposition 2 is constant across prior-information states, so the expectation and deterministic representations coincide.

We now derive the counterfactual system in proportional changes. Let $\hat{x} = x'/x$ denote the ratio of the counterfactual value of any variable $x$ to its baseline value. Start from Proposition 2 and define the inner-nest inclusive value

$$S_{od} := \sum_{s \in o} \bar{\pi}_{sd|o} c_s^{-\zeta/\lambda_{od}}. \tag{44}$$

Then the price index can be written as

$$P_d = \left( \sum_o \bar{\pi}_{od} \tau_{od}^{-\zeta/\lambda} \phi_{od}^{1/\lambda} S_{od}^{\lambda_{od}/\lambda} \right)^{-\lambda/\zeta}. \tag{45}$$

Define

$$A_{od} := \bar{\pi}_{od} \tau_{od}^{-\zeta/\lambda} \phi_{od}^{1/\lambda} S_{od}^{\lambda_{od}/\lambda}.$$

It follows that

$$\pi_{od} = \frac{A_{od}}{\sum_{o'} A_{o'd}}, \qquad P_d = \left( \sum_o A_{od} \right)^{-\lambda/\zeta}.$$

Taking proportional changes yields

$$\widehat{P}_d = \left( \sum_o \pi_{od} \widehat{A}_{od} \right)^{-\lambda/\zeta}, \tag{46}$$

and

$$\widehat{\pi}_{od} = \frac{\widehat{A}_{od}}{\sum_{o'} \pi_{o'd} \widehat{A}_{o'd}}. \tag{47}$$

To express $\widehat{A}_{od}$ in terms of primitives, first note that $\phi_{od} = I_{od}^{\gamma}$. Second, for all suppliers in origin $o$, unit costs move proportionally with the origin wage and price index:

$$\widehat{c}_s = \widehat{w}_o^{1-\alpha} \widehat{P}_o^{\alpha}.$$

Let $\widehat{\lambda}_{od} := \lambda'_{od}/\lambda_{od}$. Using the baseline inner-nest probabilities,

$$\pi_{sd|o} = \frac{\bar{\pi}_{sd|o} c_s^{-\zeta/\lambda_{od}}}{\sum_{r \in o} \bar{\pi}_{rd|o} c_r^{-\zeta/\lambda_{od}}}, \tag{48}$$

the counterfactual inner-nest inclusive value can be rewritten as

$$\left(S'_{od}\right)^{\lambda'_{od}/\lambda} = \left(\widehat{w}_o^{1-\alpha} \widehat{P}_o^{\alpha}\right)^{-\zeta/\lambda} \left(S_{od}\right)^{\lambda_{od}/\lambda} \left( \sum_{s \in o} \bar{\pi}_{sd|o}^{\,1-1/\widehat{\lambda}_{od}} \pi_{sd|o}^{\,1/\widehat{\lambda}_{od}} \right)^{\lambda'_{od}/\lambda}. \tag{49}$$

Holding $\tau_{od}$ and $\bar{\pi}_{od}$ fixed in the counterfactual then implies

$$\widehat{A}_{od} = \widehat{I}_{od}^{\gamma/\lambda} \left( \sum_{s \in o} \bar{\pi}_{sd|o}^{\,1-1/\widehat{\lambda}_{od}} \pi_{sd|o}^{\,1/\widehat{\lambda}_{od}} \right)^{\lambda'_{od}/\lambda} \left(\widehat{w}_o^{1-\alpha} \widehat{P}_o^{\alpha}\right)^{-\zeta/\lambda}. \tag{50}$$

The labor-market clearing condition

$$w_o L_o = \sum_d \pi_{od} w_d L_d$$

implies

$$\widehat{w}_o w_o L_o = \sum_d \widehat{\pi}_{od} \pi_{od} \widehat{w}_d w_d L_d. \tag{51}$$

Finally, indirect utility implies

$$\widehat{V}_d = \left( \frac{\widehat{w}_d}{\widehat{P}_d} \right)^{\beta}. \tag{52}$$

To compute welfare changes for a given fiber rollout, we take the observed shocks $\widehat{I}_{od}$ as primitives, initialize guesses for $\{\widehat{w}_o, \widehat{P}_o\}_o$, and iterate on the system given by (46)–(51) until convergence. The resulting fixed point delivers $\{\widehat{P}_d, \widehat{w}_d\}_d$, and welfare changes then follow directly from

$$\widehat{V}_d = \left( \frac{\widehat{w}_d}{\widehat{P}_d} \right)^{\beta}.$$

# D Appendix: Estimation

This appendix provides additional technical details on the estimation strategy outlined in Section 5. The discussion proceeds in three main parts:

1. **Treatment of unobservables:** We explain how we proxy for unobserved components such as supplier costs, priors, and trade frictions.

2. **Computational implementation:** We establish the equivalence between the nested logit likelihood and Poisson pseudo-maximum likelihood (PPML) estimation, which makes our large-scale estimation computationally feasible.

3. **Endogeneity correction:** We describe the control function approach that addresses potential endogeneity of fiber connectivity.

## D.1 Treatment of Unobservables

The nested logit probabilities for supplier choice, given in equations (12)–(13), depend on three structural composites we aim to estimate—$\lambda\bar{\eta}$, $\lambda\eta$, and $\gamma/\lambda$—as well as several unobservable components. These unobservables include within-province priors $\bar{\pi}_{sd|o(s)}$, supplier marginal costs $c_{st}$, across-province priors $\bar{\pi}_{od}$, and bilateral trade costs $\tau_{od}$.

Direct estimation of these high-dimensional unobservables is infeasible. Instead, we combine proxy variables with fixed effects to control for them. We now detail our approach for each type of unobservable.

**Within-Province Priors.** The within-province priors $\bar{\pi}_{sd|o(s)}$ capture baseline preferences for specific suppliers within an origin province. Including a separate fixed effect for each supplier-destination pair would be computationally prohibitive given our large dataset.

We proxy for these priors using temporal average of cost shares. Specifically, we assume that priors are correlated with the average cost share of a supplier-destination pair over time:

$$\ln \bar{\pi}_{sd|o(s)} = \bar{\kappa}_0 + \kappa_0 \ln \xi_{sd|o(s)},$$

where $\xi_{sd|o(s)}$ represents the supplier–destination cost share averaged across all years in our sample. By using time-averaged shares as proxies, we ensure that our structural parameters are identified from year-to-year variation in fiber connectivity rather than from cross-sectional differences in trading relationships.

**Supplier Marginal Costs.** Supplier marginal costs $c_{st}$ are unobserved, yet they critically affect buyer choices. We construct a proxy based on the supplier's position in the production network. Empirical evidence suggests that more productive firms attract more buyers, leading to higher network centrality (larger weighted outdegree).

Let $m_{st}$ denote the total sales (weighted outdegree) of supplier $s$ at time $t$. We normalize this by the province-level total to obtain a relative measure:

$$\widetilde{m}_{st} = \frac{m_{st}}{\sum_{s' \in o(s)} m_{s't}},$$

which captures the supplier's market share within its origin province. We then assume marginal costs follow the log-linear relationship:

$$\ln c_{st} = \delta_{ot} + \vartheta \ln \widetilde{m}_{st}.$$

*Why this normalization?* Within each origin province, buyers compare suppliers from the same location. The inner-nest estimation therefore identifies only *relative* differences in costs across suppliers within a province. Province-level cost shifters $\delta_{ot}$ drop out of within-province comparisons and are instead absorbed by fixed effects in the outer-nest estimation (where buyers choose between provinces). This normalization thus aligns our empirical specification with the economic structure of the choice problem.

Substituting these proxies into the model yields simplified inner-nest choice probabilities:

$$\pi_{sd|o(s),t} = \frac{\exp\left(\kappa_0 \ln \xi_{sd|o(s)} + \kappa_1 \ln \widetilde{m}_{st} + \kappa_2 \ln \widetilde{m}_{st} \cdot \ln I_{o(s)d,t}\right)}{\sum_{s' \in o(s)} \exp\left(\kappa_0 \ln \xi_{s'd|o(s)} + \kappa_1 \ln \widetilde{m}_{s't} + \kappa_2 \ln \widetilde{m}_{s't} \cdot \ln I_{o(s)d,t}\right)}, \tag{53}$$

where $\kappa_1 = -\zeta \vartheta \bar{\eta}$ and $\kappa_2 = -\zeta \vartheta \eta$.

**Province-Level Priors and Trade Costs.** In the outer nest, buyers choose which province to source from. This decision depends on province-level priors $\bar{\pi}_{od}$ (baseline preferences for province pairs) and bilateral trade costs $\tau_{od}$.

In our setting, these components are time-invariant: trade costs depend on factors that we assume do not change over our study period, while baseline preferences reflect long-standing relationships. We therefore include origin-destination fixed effects $\delta_{od}$ in the outer-nest estimation, which fully absorb both $\bar{\pi}_{od}$ and $\tau_{od}$. This approach ensures identification of our parameters of interest comes entirely from temporal variation in fiber connectivity rather than from cross-sectional differences in trade costs or preferences:

$$\pi_{od,t} = \frac{\exp\left(\delta_{ot} + \delta_{od} + (\gamma/\lambda)\ln I_{od,t} + \kappa_3 \widetilde{IV}_{od,t}\right)}{\sum_{o'} \exp\left(\delta_{o't} + \delta_{o'd} + (\gamma/\lambda)\ln I_{o'd,t} + \kappa_3 \widetilde{IV}_{o'd,t}\right)}, \tag{54}$$

where $\widetilde{IV}_{od,t} = (1 - (\kappa_2/\kappa_1)\ln I_{od,t}) \ln\left(\sum_{s\in o} \exp\left(\kappa_0 \ln \xi_{sd|o} + \kappa_1 \ln \widetilde{m}_{st} + \kappa_2 \ln \widetilde{m}_{st} \cdot \ln I_{od,t}\right)\right)$.

*Remark on identification.* Note that the prefactor $1 - (\kappa_2/\kappa_1)\ln I_{od,t}$ approximates $\bar{\eta}\lambda_{od}$ under a first-order expansion, so $\widetilde{IV}_{od,t} \approx \bar{\eta}\lambda_{od} \ln S_{od,t}$. This rescaling ensures that the constructed regressor depends only on the first-stage coefficients $(\kappa_0, \kappa_1, \kappa_2)$ and hence only on the ratio $\eta/\bar{\eta}$—not on the level of $\bar{\eta}$. The outer-nest coefficient $\kappa_3$ then identifies the ratio: since the model requires $\kappa_3 \cdot \widetilde{IV}_{od,t} = (\lambda_{od}/\lambda)\ln S_{od,t} \approx (1/(\lambda\bar{\eta}))\widetilde{IV}_{od,t}$, we obtain $\kappa_3 = 1/(\lambda\bar{\eta})$. Together with $\eta/\bar{\eta} = \kappa_2/\kappa_1$ from the inner nest, the ratio of all information-cost parameters to the outer-nest scale $\lambda$ is identified sequentially. Because $\lambda$ enters the outer nest as a common scale factor—absorbed into fixed effects for trade costs—it cannot be separately identified from the estimation alone and is treated as a normalization.

## D.2 Likelihood Equations and PPML Implementation

### Computational Feasibility

Our dataset contains millions of supplier-buyer-year observations. Standard nested logit estimation routines, which rely on numerical optimization of the full likelihood function, are computationally infeasible at this scale.

We overcome this challenge by exploiting a mathematical equivalence: the first-order conditions of the nested logit maximum likelihood problem coincide with those of a Poisson regression.[11] This allows us to estimate the model using Poisson pseudo-maximum likelihood (PPML), which is fast, numerically stable, and handles large datasets efficiently. Moreover, PPML can be implemented sequentially—we estimate the inner nest first, then use those estimates in the outer nest—further improving computational tractability.

[11] The equivalence between multinomial logit and Poisson regressions has a long history since Baker (1994) with recent application to international trade in Eaton et al. (2013) and high-dimensional settings in Panigrahi (2021). Here, we extend the equivalence to nested logit models.

**PPML-Nested Logit Equivalence**

We now formally establish the equivalence between nested logit maximum likelihood estimation and Poisson pseudo-maximum likelihood (PPML). The key insight is that nested logit likelihood equations can be rewritten as moment conditions equating model-predicted probabilities to empirical expenditure shares—precisely the form that PPML estimation targets.

**Proposition 3** (PPML Implementation of Nested-Logit Likelihood Equations)**.** *Model Setup. Consider a nested-logit model with inner-nest choice probabilities*

$$\pi_{sd|o(s),t} = \frac{\exp(v_{sd|o(s),t})}{\sum_{s' \in o(s)} \exp(v_{s'd|o(s),t})}, \tag{55}$$

*and outer-nest choice probabilities*

$$\pi_{od,t} = \frac{\exp(u_{od,t})}{\sum_{o'} \exp(u_{o'd,t})}, \tag{56}$$

*where the systematic utility indices are*

$$v_{sd|o(s),t} = \kappa_0 \ln \xi_{sd|o(s)} + \kappa_1 \ln \widetilde{m}_{st} + \kappa_2 \ln \widetilde{m}_{st} \cdot \ln I_{o(s)d,t}, \tag{57}$$

$$u_{od,t} = \delta_{ot} + \delta_{od} + (\gamma/\lambda) \ln I_{od,t} + \kappa_3 \widetilde{IV}_{od,t}. \tag{58}$$

*Empirical Counterparts. Define the empirical expenditure shares as*

$$Y_{sd,t} = \frac{1}{M_{d,t}} \sum_{b \in \mathcal{B}_{d,t}} \frac{\text{Cost Share}_{sbt}}{\sum_{s' \in o(s)} \text{Cost Share}_{s'bt}}, \qquad Y_{od,t} = \frac{1}{M_{d,t}} \sum_{b \in d} \sum_{s \in o} \text{Cost Share}_{sbt},$$

*where $M_{d,t}$ is the number of buyers in destination $d$ at time $t$.*

*PPML Specification. Specify the PPML models for the inner and outer nests as*

$$\mathbb{E}\left[Y_{sd,t} \mid \ln \xi_{sd|o(s)}, \ln \widetilde{m}_{st}, \ln I_{o(s)d,t}, \delta_{o(s)d,t}\right] = \exp\left(v_{sd|o(s),t} + \delta_{o(s)d,t}\right), \tag{59}$$

$$\mathbb{E}\left[Y_{od,t} \mid \ln I_{od,t}, \widetilde{IV}_{od,t}, \delta_{ot}, \delta_{od}, \delta_{dt}\right] = \exp(u_{od,t} + \delta_{d,t}), \tag{60}$$

*where $\delta_{o(s)d,t}$ is an origin-destination-time fixed effect for the inner nest, and $\delta_{d,t}$ is a destination-time fixed effect for the outer nest.*

*Result. The first-order conditions of the PPML estimators in* (59)–(60) *coincide exactly with the likelihood equations of the nested-logit model. Thus, PPML with appropriately specified fixed effects provides a computationally tractable implementation of nested-logit maximum likelihood estimation.*

*Proof.* The proof proceeds in four steps: (1) decompose the full likelihood into inner and outer components, (2) derive the inner-nest likelihood equations, (3) derive the outer-nest likelihood equations, and (4) show that PPML score conditions coincide with these likelihood equations.

**Step 1: Likelihood Decomposition.**

At time $t$, the nested-logit model assigns a buyer in destination $d$ the overall probability of choosing supplier $s$:

$$\pi_{sd,t} = \pi_{o(s)d,t} \cdot \pi_{sd|o(s),t}.$$

Observed expenditure shares Cost $\text{Share}_{sbt}$ serve as sampling weights. The log-likelihood decomposes additively:

$$\ell \propto \sum_{s,b,t} \text{Cost Share}_{sbt} \left( \ln \pi_{sd(b)|o(s),t} + \ln \pi_{o(s)d(b),t} \right). \tag{61}$$

This separability allows us to analyze each nest independently.

**Step 2: Inner-Nest Likelihood Equations.**

We focus on the inner-nest component of (61), which governs supplier choice conditional on the origin province. Because the inner-nest parameters enter only through the conditional probabilities $\pi_{sd|o(s),t}$, their score equations are identical whether we differentiate the full likelihood or the *conditional* inner-nest log-likelihood. The latter takes the form

$$\ell^{\text{cond}}_{\text{inner}} = \sum_{b,t} \sum_{o} \sum_{s \in o} \text{Cost Share}_{sbt} \ \ln \pi_{sd(b)|o,t}.$$

To re-express this in a convenient form, define the within-origin conditional expenditure weight

$$w_{sb,t} = \frac{\text{Cost Share}_{sb,t}}{\sum_{s' \in o(s)} \text{Cost Share}_{s',bt}},$$

and the origin-level expenditure total for buyer $b$, $W_{o,b,t} := \sum_{s' \in o} \text{Cost Share}_{s'bt}$. Then $\text{Cost Share}_{sbt} = w_{sb,t}\, W_{o(s),b,t}$, and substituting into the conditional log-likelihood gives

$$\ell^{\text{cond}}_{\text{inner}} = \sum_{b,t} \sum_{o} W_{o,b,t} \sum_{s \in o} w_{sb,t} \ \ln \pi_{sd(b)|o,t}.$$

Using $\ln \pi_{sd|o,t} = v_{sd|o,t} - \ln\big(\sum_{s' \in o} \exp(v_{s'd|o,t})\big)$, the second term inside the sum is constant across $s \in o$ and multiplies $\sum_{s \in o} w_{sb,t} = 1$, so

$$\ell^{\text{cond}}_{\text{inner}} \propto \sum_{d,t} \sum_{s} \left( \sum_{b \in d} W_{o(s),b,t}\, w_{sb,t} \right) v_{sd|o(s),t} - \sum_{d,t} \sum_{o} \left( \sum_{b \in d} W_{o,b,t} \right) \ln \left( \sum_{s' \in o} \exp(v_{s'd|o,t}) \right). \tag{62}$$

The PPML implementation targets the normalized dependent variable $Y_{sd,t} := \frac{1}{M_{d,t}} \sum_{b \in d} w_{sb,t}$, which conditions on origin. The PPML score conditions for this normalized variable—derived in Step 4 below—coincide with those of the conditional likelihood (62) up to a proportionality constant that does not affect the parameter estimates. Specifically, differentiating the PPML criterion with respect to $v_{sd|o(s),t}$ and setting to zero yields

$$Y_{sd,t} = \pi^{*}_{sd|o(s),t}, \tag{63}$$

so that at the optimum the model-implied conditional probability equals the average empirical conditional expenditure share across all buyers in destination $d$.

**Step 3: Outer-Nest Likelihood Equations.**

Define the aggregate expenditure from destination $d$ to origin $o$:

$$W_{od,t} = \sum_{b \in d} \sum_{s \in o} \text{Cost Share}_{sbt}.$$

Since each buyer's expenditure shares sum to unity, we have the adding-up constraint:

$$\sum_{o} W_{od,t} = M_{d,t}.$$

Substituting (56) into the full likelihood yields the outer-nest likelihood:

$$\ell_{\text{outer}} \propto \sum_{d,t} \sum_{o} W_{od,t}\, u_{od,t} - \sum_{d,t} M_{d,t} \ln\left( \sum_{o'} \exp(u_{o'd,t}) \right). \tag{64}$$

Differentiating with respect to $u_{od,t}$ and setting to zero:

$$\frac{\partial \ell_{\text{outer}}}{\partial u_{od,t}} = W_{od,t} - M_{d,t}\, \pi_{od,t} = 0.$$

Rearranging yields the outer-nest likelihood equation:

$$\pi^*_{od,t} = \frac{W_{od,t}}{M_{d,t}} = Y_{od,t}. \tag{65}$$

At the maximum likelihood estimate, the model-implied probability of selecting origin $o$ equals the share of total expenditure in destination $d$ accounted for by suppliers from origin $o$.

Equations (63) and (65) establish that nested-logit MLE equates model-predicted choice probabilities to empirical expenditure shares. We now show that PPML produces identical conditions.

**Step 4: PPML Score Conditions.**

*Inner nest.* Under the PPML specification (59), the fixed effect $\delta_{o(s),d,t}$ generates the score condition:

$$\sum_{s \in o} (Y_{sd,t} - \widehat{Y}_{sd,t}) = 0.$$

By construction of the dependent variable, $\sum_{s \in o} Y_{sd,t} = 1$ for each origin-destination-time triplet $(o, d, t)$. The score condition therefore implies $\sum_{s \in o} \hat{Y}_{sd,t} = 1$.

Since $\hat{Y}_{sd,t} = \exp(v_{sd|o,t} + \delta_{od,t})$, this normalization constraint yields:

$$\widehat{Y}_{sd,t} = \frac{\exp(v_{sd|o(s),t})}{\sum_{s' \in o} \exp(v_{s'd|o(s),t})} = \pi_{sd|o(s),t}.$$

Thus, the PPML fitted values coincide with the inner-nest choice probabilities, and the score conditions for $(\kappa_0, \kappa_1, \kappa_2)$ are identical to the likelihood equation (63).

*Outer nest.* Under the PPML specification (60), the destination-time fixed effect $\delta_{d,t}$ generates the score condition:

$$\sum_o (Y_{od,t} - \hat{Y}_{od,t}) = 0.$$

Since $\sum_o Y_{od,t} = 1$ by construction, this implies $\sum_o \hat{Y}_{od,t} = 1$, yielding:

$$\hat{Y}_{od,t} = \frac{\exp(u_{od,t})}{\sum_{o'} \exp(u_{o'd,t})} = \pi_{od,t}.$$

The PPML score equations for $(\gamma/\lambda, \kappa_3)$ therefore coincide with the outer-nest likelihood equation (65).

□

For both nests, the PPML score conditions impose exactly the same moment conditions—equating predicted probabilities to empirical shares—as the nested-logit likelihood equations. The fixed effects ensure that predicted probabilities satisfy the required adding-up constraints. Therefore, PPML with the specified fixed effects provides a computationally tractable implementation of nested-logit maximum likelihood estimation.

## D.3 Control Function Approach

### The Endogeneity Problem

Fiber connectivity $I_{od,t}$ may be endogenous. If unobserved determinants of trade, captured in the error terms $\varepsilon_{sd,t}$ and $\varepsilon_{od,t}$, are correlated with fiber deployment, our estimates of the structural parameters $\gamma/\lambda$, $\lambda\bar{\eta}$, and $\lambda\eta$ would be biased.

### Identification Strategy

We address this concern using a control function approach (Wooldridge, 2015) based on our instrument $Z_{od,t}$, described in Section 3. The control function method proceeds in two steps:

1. **First stage:** Regress fiber connectivity on the instrument to decompose it into an exogenous predicted component and a residual that captures endogenous variation.

2. **Second stage:** Include these residuals directly in the structural estimation, thereby "purging" the endogeneity from the fiber connectivity variable.

Formally, suppose the PPML models contain unobserved error terms:

$$Y_{sd,t} = \exp\left(v_{sd|o(s),t} + \delta_{o(s),d,t} + \varepsilon_{sd,t}\right), \tag{66}$$

$$Y_{od,t} = \exp(u_{od,t} + \delta_{d,t} + \varepsilon_{od,t}). \tag{67}$$

For identification, we require five conditions. The first ensures the instrument is informative; the second characterizes the source of endogeneity; the third and fourth establish exogeneity of the control function and instrument; and the fifth guarantees unique identification of the parameters.

**Assumption 1** (Relevance). *The instrument $Z_{od,t}$ is correlated with fiber connectivity:*

$$Cov\left[Z_{od,t}, \ln I_{od,t}\right] \neq 0.$$

*Let*

$$\ln I_{od,t} = \alpha_0 + \alpha_1 Z_{od,t} + \nu_{od,t}$$

*be the reduced form, where $\widehat{\ln I_{od,t}}$ is the fitted value (the exogenous component) and $\widehat{\nu}_{od,t}$ is the residual (the potentially endogenous component).*

**Assumption 2** (Structural Endogeneity). *Endogeneity operates through the first-stage residual:*

$$\begin{aligned}\varepsilon_{sd,t} &= h_1\left(\nu_{od,t}\right)\ln \widetilde{m}_{st} + \mu_{sd,t},\\ \varepsilon_{od,t} &= h_2\left(\nu_{od,t}\right) + \mu_{od,t},\end{aligned}$$

*where $h_1(\cdot)$ and $h_2(\cdot)$ are unknown measurable functions, and $\mu_{sd,t}$, $\mu_{od,t}$ are idiosyncratic errors uncorrelated with the instrument. In the inner nest, the endogenous component interacts with the supplier characteristic $\ln \widetilde{m}_{st}$ because the inner-nest specification includes origin–destination–year fixed effects $\delta_{o(s)d,t}$, which absorb any additive function of $\nu_{od,t}$ alone. The endogeneity concern is therefore specifically about the interaction regressor $\ln \widetilde{m}_{st} \cdot \ln I_{od,t}$: unobserved connectivity shocks may differentially affect trade flows through suppliers of different sizes. In the outer nest, which conditions on origin–year, destination–year, and pair fixed effects separately, $h_2(\nu_{od,t})$ retains time-varying within-pair variation and is not absorbed.*

**Assumption 3** (Control Function Exogeneity). *Conditional on the first-stage residual, the remaining error is mean-independent of the regressors:*

$$\begin{aligned}\mathbb{E}\left[\mu_{sd,t} \mid Z_{od,t}, \widehat{\ln I_{od,t}}, \ln \xi_{sd|o(s)}, \ln \widetilde{m}_{st}, \delta_{o(s)d,t}\right] &= 0\\ \mathbb{E}\left[\mu_{od,t} \mid Z_{od,t}, \widehat{\ln I_{od,t}}, \widetilde{\widetilde{IV}}_{od,t}, \delta_{od}, \delta_{ot}, \delta_{dt}\right] &= 0\end{aligned}$$

*This is the key identifying assumption: once we control for the first-stage residual, the remaining variation in fiber connectivity is exogenous.*

**Assumption 4** (Exclusion Restrictions). *The instrument affects trade only through fiber connectivity:*

$$\begin{aligned}\mathbb{E}\left[Z_{od,t}\varepsilon_{sd,t} \mid \ln I_{od,t}, \ln \xi_{sd|o(s)}, \ln \widetilde{m}_{st}, \delta_{o(s)d,t}\right] &= 0\\ \mathbb{E}\left[Z_{od,t}\varepsilon_{od,t} \mid \ln I_{od,t}, \delta_{od}, \delta_{ot}, \delta_{dt}\right] &= 0\end{aligned}$$

*This standard exclusion restriction rules out direct effects of the instrument on trade outcomes.*

**Assumption 5** (Rank Condition)**.** *The model is locally identified: The Jacobian of the score conditions with respect to the parameter vector $\theta \equiv (\kappa_0, \kappa_1, \kappa_2, \kappa_3, \gamma/\lambda)$, evaluated at the true parameter values, has full rank:*

$$rank\left(\mathbb{E}\left[\frac{\partial g(\theta)}{\partial \theta'}\right]\right) = dim\left(\theta\right),$$

*where the score vector $g(\theta)$ collects the first-order conditions:*

$$\begin{aligned}
g(\theta) &= \left[g_{\kappa_0}(\theta), g_{\kappa_1}(\theta), g_{\kappa_2}(\theta), g_{\kappa_3}(\theta), g_{\gamma/\lambda}(\theta)\right]', \\
g_{\kappa_0}(\theta) &= \sum_{s,d,t} \ln \xi_{sd|o(s)} \left[Y_{sd,t} - \exp\left(v_{sd|o(s),t} + \delta_{o(s),d,t}\right)\right], \\
g_{\kappa_1}(\theta) &= \sum_{s,d,t} \ln \tilde{m}_{st} \left[Y_{sd,t} - \exp\left(v_{sd|o(s),t} + \delta_{o(s),d,t}\right)\right], \\
g_{\kappa_2}(\theta) &= \sum_{s,d,t} \left(\ln \tilde{m}_{st} \cdot \ln I_{od,t}\right) \left[Y_{sd,t} - \exp\left(v_{sd|o(s),t} + \delta_{o(s),d,t}\right)\right], \\
g_{\kappa_3}(\theta) &= \sum_{o,d,t} \widetilde{\widetilde{IV}}_{od,t} \left[Y_{od,t} - \exp\left(u_{od,t} + \delta_{d,t}\right)\right], \\
g_{\gamma/\lambda}(\theta) &= \sum_{o,d,t} \ln I_{od,t} \left[Y_{od,t} - \exp\left(u_{od,t} + \delta_{d,t}\right)\right].
\end{aligned}$$

*This ensures that the system of moment conditions has a unique solution in a neighborhood of the true parameter values.*

**Identification Result**

**Lemma** (Identification via Control Function)**.** *Under Assumptions 1–5, the PPML estimators based on the augmented (control-function) inner- and outer-nest moment conditions consistently identify the structural parameters $(\kappa_0, \kappa_1, \kappa_2, \kappa_3, \gamma/\lambda)$.*

*Proof.* The proof proceeds by showing that augmenting the PPML specifications with first-stage residuals eliminates the endogeneity bias.

**Step 1: First-stage decomposition.** From the reduced-form equation for fiber connectivity,

$$\ln I_{od,t} = \alpha_0 + \alpha_1 Z_{od,t} + \nu_{od,t},$$

we obtain fitted values $\widehat{\ln I_{od,t}}$ and first-stage residuals $\widehat{\nu}_{od,t} = \ln I_{od,t} - \widehat{\ln I_{od,t}}$. By Assumption 1, the fitted value $\widehat{\ln I_{od,t}}$ is determined solely by the exogenous instrument $Z_{od,t}$. By Assumption 2, any correlation between fiber connectivity and the structural errors operates exclusively through $\nu_{od,t}$.

**Step 2: Augmented inner-nest specification.** We augment the inner-nest PPML model by including the interaction $\ln \widetilde{m}_{st} \cdot \widehat{\nu}_{od,t}$:

$$Y_{sd,t} = \exp\left(\kappa_0 \ln \xi_{sd|o(s)} + \kappa_1 \ln \widetilde{m}_{st} + \kappa_2 \ln \widetilde{m}_{st} \cdot \ln I_{o(s)d,t} + v_0 \ln \widetilde{m}_{st} \cdot \widehat{\nu}_{od,t} + \delta_{o(s)d,t} + \varepsilon_{sd,t}\right).$$

The coefficient $v_0$ on the control function term captures the endogenous component of the fiber-supplier interaction.

**Step 3: Augmented outer-nest specification.** For the outer nest, we make two modifications. First, we construct a "purged" inclusive value $\widetilde{\widetilde{IV}}_{od,t}$ using only the exogenous component of fiber connectivity:

$$\widetilde{\widetilde{IV}}_{od,t} = \left(1 - (\kappa_2/\kappa_1)\,\widehat{\ln I_{od,t}}\right) \ln\left(\sum_{s\in o} \exp\left(\kappa_0 \ln \xi_{sd|o} + \kappa_1 \ln \widetilde{m}_{st} + \kappa_2 \ln \widetilde{m}_{st} \cdot \widehat{\ln I_{od,t}}\right)\right).$$

Second, we include the first-stage residual directly as a control:

$$Y_{od,t} = \exp\left(\delta_{ot} + \delta_{od} + (\gamma/\lambda)\ln I_{od,t} + \kappa_3 \widetilde{\widetilde{IV}}_{od,t} + v_1 \widehat{\nu}_{od,t} + \delta_{dt} + \varepsilon_{od,t}\right).$$

The coefficient $v_1$ captures the endogenous component of fiber connectivity in the province-level choice.

**Step 4: Conditional exogeneity.** By Assumption 2, the inner-nest structural error decomposes as $\varepsilon_{sd,t} = h_1(\nu_{od,t}) \ln \widetilde{m}_{st} + \mu_{sd,t}$. Any additive function of $\nu_{od,t}$ alone is absorbed by the origin–destination–year fixed effect $\delta_{o(s)d,t}$, so the endogeneity concern resides entirely in the interaction term. The control function $v_0 \ln \widetilde{m}_{st} \cdot \widehat{\nu}_{od,t}$ absorbs $h_1(\nu_{od,t}) \ln \widetilde{m}_{st}$, the supplier-varying endogenous component (under mild regularity conditions on $h_1$). For the outer nest, $\varepsilon_{od,t} = h_2(\nu_{od,t}) + \mu_{od,t}$, and the control function $v_1 \widehat{\nu}_{od,t}$ absorbs $h_2(\nu_{od,t})$ directly. By Assumption 3, the remaining idiosyncratic errors satisfy:

$$\begin{aligned} \mathbb{E}\left[\mu_{sd,t} \mid \ln \xi_{sd|o(s)}, \ln \widetilde{m}_{st}, \ln I_{od,t}, \widehat{\nu}_{od,t}, \delta_{o(s)d,t}\right] &= 0, \\ \mathbb{E}\left[\mu_{od,t} \mid \ln I_{od,t}, \widetilde{\widetilde{IV}}_{od,t}, \widehat{\nu}_{od,t}, \delta_{ot}, \delta_{od}, \delta_{dt}\right] &= 0. \end{aligned}$$

Thus, both augmented PPML models satisfy the conditional moment restrictions required for consistent estimation of the structural parameters.

**Step 5: Identification.** Assumption 4 (exclusion) ensures that the instrument affects structural utilities only through $\ln I_{od,t}$, so variation in $\widehat{\ln I_{od,t}}$ provides valid exogenous variation. Assumption 5 (rank) guarantees that the mapping from parameters to PPML score conditions is locally invertible. Together, these conditions ensure that the augmented PPML first-order conditions have a unique solution at the true parameter values. □

#### Implementation Summary

In practice, we implement the control function approach as follows:

1. Regress $\ln I_{od,t}$ on the instrument $Z_{od,t}$ and fixed effects to obtain residuals $\widehat{\nu}_{od,t}$.
2. In the inner nest, include both $\ln I_{od,t}$ and the interaction $\ln \widetilde{m}_{st} \cdot \widehat{\nu}_{od,t}$ as regressors.

3. In the outer nest, construct a "purged" inclusive value $\widetilde{\widetilde{IV}}_{od,t}$ using only the exogenous component $\widehat{\ln I_{od,t}}$, and include both $\ln I_{od,t}$ and $\widehat{\nu}_{od,t}$ as regressors.

By conditioning on the residuals $\widehat{\nu}_{od,t}$, we effectively control for the endogenous variation in fiber connectivity, allowing us to recover unbiased estimates of the structural parameters.

## D.4 Estimation Results

Table A14 presents results from our PPML estimation with control functions. The table reports estimates in two panels. In the inner nest, the key coefficients are the seller–year proxy ($\kappa_1$), and their interaction with fiber connectivity ($\kappa_2$). In the outer nest, the key coefficients are the direct effect of fiber connectivity ($\gamma/\lambda$) and the inner nest inclusive value ($\kappa_3$).

Table A14: **Estimation of Information Acquisition and Communication Costs**

| Parameter | | Value |
|---|---|---|
| **Inner Nest** | | |
| Seller×Destination Proxy | $\kappa_0$ | 0.844*** |
| | | (0.00291) |
| Seller×Year Proxy | $\kappa_1$ | 0.183*** |
| | | (0.00224) |
| Fiber Connectivity × Seller×Year Proxy | $\kappa_2$ | 0.00656*** |
| | | (0.00232) |
| Control Function | $\nu_0$ | 0.00678 |
| | | (0.0143) |
| Pseudo-$R^2$ | | 0.684 |
| Observations | | 5,089,428 |
| **Outer Nest** | | |
| Fibre Connectivity | $\gamma/\lambda$ | 0.462*** |
| | | (0.110) |
| Inner Nest Inclusive Value | $\kappa_3$ | 1.501*** |
| | | (0.0461) |
| Control Function | $\nu_1$ | -0.970*** |
| | | (0.381) |
| **Fixed Effects:** | | |
| Origin×Year | | ✓ |
| Destination×Year | | ✓ |
| Origin×Destination | | ✓ |
| Pseudo-$R^2$ | | 0.989 |
| Observations | | 43,465 |

**Note**: In the inner nest, each observation pertains to a supplier firm, a destination province, and a year. In the outer nest, each observation pertains to an origin province, a destination province, and a year. To address endogeneity concerns, estimation is done through 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.

## D.5 Recovering Structural Parameters

The final step translates our reduced-form PPML coefficients into the structural composites of interest. The nested logit structure with information frictions implies the following mapping:

$$\lambda\bar{\eta} = \frac{1}{\kappa_3}, \quad \lambda\eta = \frac{\kappa_2}{\kappa_1\kappa_3}, \quad \gamma/\lambda \text{ is directly estimated.}$$

These formulas allow us to translate the PPML estimates in Table A14 into interpretable structural elasticities that quantify how improvements in fiber connectivity affect communication costs and information acquisition costs. Note that the outer-nest entropy weight $\lambda$ enters only as a common scale; the composites $\lambda\bar{\eta}$, $\lambda\eta$, and $\gamma/\lambda$—which govern all observable choice probabilities—are fully identified.

## D.6 Goodness of Fit

With the structural parameters in hand, we assess whether the calibrated model can replicate observed patterns in the data before using it for the counterfactual analysis in Section 6. The structural estimation uses observed trade shares to estimate the model parameters, and the goodness-of-fit test checks whether the model replicates *changes* in trade shares. The model is estimated from levels, while the test evaluates its ability to predict temporal changes, providing an independent assessment of the model's empirical content. Following Adão et al. (2025), we regress observed outcomes on model-implied outcomes using average period changes: $\Delta^{2012-2019}y_{od} = \alpha + \beta_{\text{fit}}\Delta^{2012-2019}\widehat{y}_{od} + \varepsilon_{od}$. A slope of $\beta_{\text{fit}} = 1$ indicates that predicted and observed values move one-for-one.

Table A15 reports the results for cost shares and cost share HHI. The model exhibits a close fit to the data: for cost shares, we obtain $\hat{\beta}_{\text{fit}} = 1.014$ (s.e. 0.029), and for cost share HHI, $\hat{\beta}_{\text{fit}} = 1.057$ (s.e. 0.248). In both cases, we fail to reject the null hypothesis of one-for-one correspondence. The associated $p$-values of 0.630 and 0.818 for cost shares and cost share HHI, respectively, confirm that the calibrated model reproduces both the direction and magnitude of changes in how firms allocate spending across suppliers.

Table A15: **Goodness of Fit Tests**

| | Cost Share (1) | Cost Share HHI (2) |
|---|---|---|
| Predicted Outcome | 1.014*** | 1.057*** |
| | (0.0293) | (0.2480) |
| $p$-value ($H_0 : \beta_{\text{fit}} = 1$) | [0.6304] | [0.8179] |
| Observations | 6561 | 6561 |
| First-stage $F$-statistic | 161.7 | 161.7 |

**Note**: The table reports goodness-of-fit regressions of observed outcomes on model-predicted outcomes. Standard errors are in parentheses. Brackets report the $p$-value for testing $H_0 : \beta_{\text{fit}} = 1$. Significance stars refer to the conventional test of $H_0 : \beta_{\text{fit}} = 0$. * 10%, ** 5%, *** 1% significance levels.

## Appendix References


ADÃO, R., A. COSTINOT, AND D. DONALDSON (2025): "Putting Quantitative Models to the Test: An Application to the U.S.-China Trade War," *The Quarterly Journal of Economics*, 140, 1471–1524.

BAKER, S. G. (1994): "The Multinomial-Poisson Transformation," *Journal of the Royal Statistical Society. Series D (The Statistician)*, 43, 495–504.

EATON, J., S. KORTUM, AND S. SOTELO (2013): "International Trade: Linking Micro and Macro," in *Advances in Economics and Econometrics: Tenth World Congress: Volume 2: Applied Economics*, ed. by D. Acemoglu, E. Dekel, and M. Arellano, Cambridge: Cambridge University Press, vol. 2 of *Econometric Society Monographs*, 329–370.

PANIGRAHI, P. (2021): "Endogenous Spatial Production Networks: Quantitative Implications for Trade and Productivity," *CESifo Working Paper Series*, number: 9466.

WOOLDRIDGE, J. M. (2015): "Control Function Methods in Applied Econometrics," *Journal of Human Resources*, 50, 420–445.