Communication Costs, Direct Flights and International Trade*

Taiji Furusawa Chang Sun Heiwai Tang Jiaxu Zhang †

Aug, 2026

Abstract

We build a general equilibrium model of endogenous communication, quality control, and trade. The model yields a structural gravity equation in which exogenous communication costs raise the probability of product failure and reduce trade more for products with a lower elasticity of substitution. We treat direct flight connectedness as a key determinant of communication costs and estimate its causal effect on trade using an instrument based on the discontinuity of direct flights at around 6,000 miles created by air travel regulations. Novel data on international business lawsuits show that air connectedness lowers the probability of product failure. Combining the estimates with the model, we quantify the aggregate impact of air connectedness in three episodes: the long-run improvement in air connectedness, the disruption of international travel during the COVID-19 pandemic, and the closure of Russian and Ukrainian airspace after the outbreak of the war in Ukraine.

Keywords: international direct flights; elasticity of substitution; face-to-face communication; quality control

JEL Codes: F1, R4

*We thank Bingjing Li, Haishi (Harry) Li and seminar and conference participants at HKU, NTU, CUHK Shenzhen, INSEAD, LMU Munich and the Ifo Institute, the University of Tokyo, the Asia-Pacific Trade Seminars, the TEMPEST workshop, the Korean Empirical Trade Conference and the UIBE GVC Conference for their helpful comments. We are grateful to Johannes Boehm for sharing detailed procedures of obtaining the Lexis data. We thank Wei Dai, Yifan Wang and Yijie Hao for excellent research assistance. Financial support from HKGRF (project code: 17503221) and HKU Business School (Futian, grant number SZRI2025-SPRP-03) is greatly appreciated.

†Taiji Furusawa: Graduate School of Economics, University of Tokyo, furusawa@e.u-tokyo.ac.jp. Chang Sun: HKU Business School, University of Hong Kong, sunc@hku.hk. Heiwai Tang: HKU Business School, University of Hong Kong, tangecon@hku.hk. Jiaxu Zhang: HKU Business School, University of Hong Kong, zhjiaxu@connect.hku.hk.

1 Introduction

Trade costs are crucial for understanding the patterns of trade and the welfare gains from trade, and economists have long been interested in measuring these costs. As Anderson and Van Wincoop (2004) point out, the trade costs required to match the aggregate trade data are very large (a tax equivalent of 170%), whereas the typical "trade costs" such as transportation and tariff barriers are relatively low (Hummels, 1999; Harrigan, 1993). One potential explanation for the unmeasured trade costs are communication costs. Communication of non-codifiable information is necessary for successful business partnerships, and technology and infrastructures that facilitate such communication may promote trade. In this paper, we evaluate the impact of direct flights on international trade and shed light on the importance of communication costs as a trade barrier.

Why does communication matter for trade, and why is in-person communication important in international business? Business practices offer a telling illustration: in-person contact is central to quality control in buyer–supplier relationships. For instance, Toyota describes its practice of sending executives to see for themselves how suppliers work as *genchi genbutsu* ("actual location and actual parts"). When Toyota found a quality problem with wire harnesses in 1997, the supplier's president flew to Toyota's Kentucky plant and spent time on the shop floor observing how workers assembled the harnesses before presenting countermeasures (Liker and Choi, 2004). Another example is Boeing's outsourcing of production of parts and components that went into its famous 787 aircraft. To tackle recurring quality issues in the components produced from subcontractors, Boeing sent hundreds of its engineers to supplier sites worldwide to resolve production problems in person (Tang, Zimmerman and Nelson, 2009; Gates, 2013). In both episodes, product failures emerged when buyers and suppliers did not communicate regularly in person—precisely the friction we study in this paper.

Motivated by these observations, we incorporate endogenous communication and quality control in Furusawa, Inui, Ito and Tang (2017) into Eaton and Kortum (2002) and outline a general equilibrium model to characterize how communication costs influence trade flows in different sectors. In our model, there is a certain probability that the seller's product cannot meet the buyer's standard. The buyer chooses the optimal probability of meeting her requirements, which the seller attains through costly communication with the buyer. On the other hand, exogenous communication costs, such as the difficulty of business travel, increase the total communication costs for any given probability of product failures. We then derive the optimal communication efforts and a gravity equation of trade flows involving an interaction term between the exogenous communication costs and the inverse elasticity of

substitution. Our model predicts that a reduction in the exogenous communication costs will increase trade flows. Such effects are stronger for products with a lower input elasticity of substitution. Intuitively, product failures render the seller's product useless in the buyer's production, and the costs of such failures are higher when the elasticity of substitution is lower.

The model treats the exogenous communication costs as a primitive without specifying their content. As the Toyota and Boeing episodes above illustrate, quality control relies on in-person contact, and travelling to meet a foreign supplier or buyer is much easier when the two countries are connected by direct flights. In our empirical analysis, we therefore focus on the air connectedness between countries due to direct flights as a key factor in the exogenous communication costs. We first construct a connection index for country pairs based on international direct flight data provided by the International Civil Aviation Organization (ICAO). Using this dataset, we determine whether any two airports are connected by direct flights. We then divide each country into $1^\circ \times 1^\circ$ grid cells. For any pair of grids in two countries, we consider them as "connected" if they both fall within a certain radius of connected airports. We then calculate the connection index as the share of grid pairs that are connected, using the product of population in the two grids as weights.

Our structural gravity equation predicts that the interaction of country-pair-level air connectedness and a composite elasticity measure—the product of the typical trade elasticity and the inverse elasticity of substitution in production—will positively affect trade flows. However, an Ordinary Least Squares (OLS) regression may be biased because unobserved factors that affect trade may also affect the establishment of direct flights between two countries. We overcome this challenge using an instrumental variable (IV) approach. In particular, we follow Campante and Yanagizawa-Drott (2018) and use the discontinuous drop in direct flights at approximately 6,000 miles to construct the IV. The regulations in the United States and Europe limit the hours of operation for flight crew members within a 24-hour period and require additional crew for any flight exceeding 12 hours. Combined with the arrival of long-haul aircraft in 1989, we observe the discontinuity mentioned above in later years. Based on this fact, for any two countries, we focus on grid pairs with a distance in the range of 5,500 to 6,500 miles, and calculate the population-weighted share of grid pairs with a distance between 5,500 and 6,000 miles. Although the overall connectedness between two countries may be heavily affected by their distance or other factors driving international trade, it is plausibly random whether the distance between two grids is below 6,000 miles, conditional on the distance being within the [5500, 6500] window. We use this share as our IV and find that it positively predicts the air connectedness between two countries.

We then estimate the structural gravity equation on a cross-section of trade flows and

air connectedness, using both OLS and IV regressions: conditional on origin-sector and destination-sector fixed effects and the distance controls, the effect of direct flights is identified from variation across country pairs. We find positive and statistically significant coefficients of the interaction term between air connectedness and the composite elasticity, with the IV estimates being slightly larger. Based on the preferred IV estimate, a ten percentage point increase in air connectedness will lead to a 2.5% increase in trade of a product with the median value of the composite elasticity. The effect is larger for products with a higher composite elasticity (lower elasticity of substitution in production). Products with a composite elasticity at the 75th percentile will see a 4.1% increase in trade when air connectedness increases by ten percentage points.

We show that our structural gravity estimates are robust to a battery of robustness checks. Our primary concern is that our instrumental variable, based on the discontinuity of direct flights between grid pairs within the range of 5,500 and 6,500 miles, remains correlated with unobservable factors that determine trade volumes between countries. In one of our robustness checks, we explicitly control for trade flows in 1989, a period before the discontinuity occurs. Suppose the unobservable factors are persistent over time, historical trade flows can well capture such factors. Unsurprisingly, historical trade flows strongly predict current trade flows; however, they do not diminish our IV estimates of direct flight connectedness. We address other concerns and demonstrate that our main results are robust to excluding products primarily shipped by air, using alternative trade elasticities when constructing the composite elasticity, modified versions of the IV, and various methods for constructing air-connectedness.

Next, we provide direct evidence on the key mechanism of our model: direct flights, which lead to lower communication costs, result in a lower probability of product failure. To do this, we construct grid-pair level measures of business activities and product defect events. We begin with the ORBIS database, which contains information about firms and their global ultimate owners (GUOs). It also contains the location (city) of each firm, which we geocode to identify their corresponding grids. To measure international business activities between two grids, we count the number of foreign ownership links from the GUO in one grid to the subsidiary in another grid. In general, it is difficult to obtain data on product failures. We obtain a novel measure combining ORBIS with the LexisNexis database, which allows us to identify firms involved in business lawsuits related to "defective product", following similar procedures as in Boehm (2022). We obtain a total of 559 defective-product cases between firms, spanning 437 country pairs.[1]

[1]One potential data source for product failures is consumer product recalls, as in Basker and Kamal (2021), but their data only covers imports by the United States. Grosz and Raval (2022) use data on

One benefit of using self-constructed grid-pair-level data is that we can utilize the discontinuity in direct flight connectedness at the 6,000-mile threshold and apply a regression discontinuity design (RDD), without resorting to the country-pair level IV constructed to analyze trade flows earlier. As in the trade analysis, the variation we exploit is cross-sectional, now across grid pairs on either side of the 6,000-mile threshold. We first confirm that having a distance of less than 6,000 miles significantly increases the probability of being connected by direct flights between two grids. Next, we find that the number of defective-product lawsuits does not change when direct flights connect the two grids. In contrast, the number of foreign ownership links increases significantly. We complement the grid-pair analysis with country-pair level evidence on the probability of product defects: the ratio of defective-product cases to foreign ownership links declines with a country pair's exposure to the 6,000-mile discontinuity. Using business lawsuits as a proxy for product failure and foreign ownership links as a proxy for international business activities, our results are consistent with our model's prediction that direct flights reduce communication costs, and thus, the probability of product failures.

Finally, we close our model in general equilibrium and conduct counterfactual analysis applying the "exact hat algebra" in Dekle, Eaton and Kortum (2007). We use the estimates from the structural gravity equation to inform the impact of direct flights on the exogenous communication costs. We parameterize the baseline equilibrium to the World Input-Output Database in 2002 (Timmer, Dietzenbacher, Los, Stehrer and de Vries, 2015). During the period from 2002 to 2011, most countries in our sample (31 of 39) experienced an improvement in air connectedness with the other countries. Correspondingly, our model predicts that these countries will have higher real wages as a result of the improvement. Small countries with large improvements in air connectedness gain the most: real wages rise by 1.42% in Malta and 0.89% in Slovenia. Benchmarking the model against the observed changes in the WIOD over the same window, the improvement in air connectedness explains 4.9% of the observed increase in the world's foreign share of absorption and 2.1% of the implied welfare gains, and considerably more in the countries with the largest improvements in connectedness. In a second counterfactual, we use the decline in air connectedness from 2019 to 2020 caused by the COVID-19 pandemic and find that small open economies with large declines, such as Ireland and Malta, experience the largest drops in real wages (0.34% each).

Our third counterfactual evaluates the closure of Russian and Ukrainian airspace after February 2022. The closure lowers air connectedness through two channels: direct flights

consumer complaints collected by various agencies in the US and Australia. We prefer business lawsuits because we can identify the companies involved in each case and their locations. Our data is at a global scale, and the analysis can be done at a sub-national level.

between Russia and the banning countries stop altogether, and flights between third countries whose original paths crossed the closed airspace must take longer detour routes, which we compute with a shortest-path algorithm on a global grid. Feeding both channels into the model, we find that the connection index across country pairs falls on average by about 1.8% of its 2019 level, with the decline heavily concentrated in Russia. Russia loses the most, with real wages falling by 0.14%, its expenditure shifting toward domestic goods, and its remaining foreign trade reallocating toward the countries that keep their direct flights. Among the banning countries, small open economies that rely on Russian inputs, such as Lithuania, bear the largest losses. A caveat when interpreting the quantitative results is that we only identify a particular mechanism in which the improvement or worsening of air connectedness affects trade, and the overall effect of air connectedness may be larger due to other mechanisms.

1.1 Related Literature

Our paper makes contributions to three strands of literature. First, it helps us better understand the barriers to doing business between sellers and producers. Existing studies have focused on search and contractual frictions. Bernard, Moxnes and Saito (2018) build a model in which the friction is a fixed, per-location search cost; the object of choice is the set of supplier locations a firm searches. A fall in the search cost raises sales, with the effect strongest in input-intensive industries, which they identify from the opening of a high-speed train line in Japan. Startz (2025) models both a product-search friction and a contracting friction in which opportunistic suppliers may fail to deliver as promised; in her setting Nigerian traders resolve these either by paying a fixed cost to travel and transact on the spot or by ordering remotely under relational price incentives, and she estimates the model on original survey data. Our model differs in the key friction and the resulting comparative statics. Rather than focusing on search frictions or moral hazard that can be resolved by relational contracts, we let a bilateral communication cost govern a monitoring technology that sets the continuous probability that a given transaction succeeds—an intensive margin internal to every transaction. The monitoring technology itself is adopted from Furusawa et al. (2017), who use endogenous buyer–seller communication to study how firms' offshoring reshapes their domestic production networks; we adapt it from their firm-level, industry-equilibrium setting into a multi-country Eaton and Kortum (2002) framework. We embed this object in a structural gravity equation, so the effect of communication costs is mediated by the elasticity of substitution.[2] Closing the model in general equilibrium lets us quantify the

[2]In a different context and focusing on a different mechanism, Gumpert (2018) examines the impact of communication costs in knowledge transmission within multinational firms.

impact of communication-cost shocks on trade and welfare. We also provide direct evidence that direct flights reduce the probability of product failure, a key mechanism in our model.

Second, we contribute to the literature evaluating the impact of air travel on international business.[3] Existing research provides reduced-form evidence that domestic or international air travel can boost FDI and exports between locations (Chen and Lin, 2020; Poole, 2009; Cristea, 2011; Wang and Fu, 2022). To explore the mechanisms of the positive impact, previous research tries to show heterogeneous treatment effects due to product characteristics, such as the degree of product differentiation constructed by Rauch (1999), contract intensity constructed by Nunn (2007) and sectors' R&D intensities, but such heterogeneous effects lack structural interpretations. We go beyond the reduced-form approach and provide a model of endogenous communication and product quality. The model generates a structural gravity equation that implies the impact of the exogenous communication costs on trade flows depends on the elasticity of substitution. The elasticity parameter has no bearing on the major predictions of the Eaton and Kortum (2002) model, other than affecting the level of the price index in all countries. However, it plays a crucial role in our model, as it distinguishes the impact of communication costs from conventional gravity forces. Moreover, we estimate the coefficients using exogenous variation caused by air travel regulations and aircraft technologies. To this end, we complement the studies by Söderlund (2023) and Besedeš, Chu and Murshid (2024), who use changes in airspaces due to geopolitical frictions or conflicts to causally identify the impact of air travel time on international trade, although their approaches remain reduced-form.

Third, this paper also contributes to understanding the impact of regulations on travel costs and other outcomes. For example, restrictions on airspace due to political reasons can limit the routes that flights travel, increase the costs of travel and reduce international businesses (Yilmazkuday and Yilmazkuday, 2017; Söderlund, 2023). Blonigen and Cristea (2015) study the impact of the 1978 Airline Deregulation Act and estimate the effects of airline traffic on American cities. Campante and Yanagizawa-Drott (2018) show that the combination of air travel regulations and aircraft technologies causes a discontinuous drop in direct flights at around 6,000-mile distance, which further causes some cities to have better air connections than others and to experience faster economic growth. We use the same discontinuity as in Campante and Yanagizawa-Drott (2018) for identification, but we combine it with our general equilibrium model to evaluate the effects of direct flights on trade, production and real wages.

[3]More broadly, the estimates in our paper also inform the impact of infrastructure on trade. Earlier studies have examined the impact of highways (Coşar and Demir, 2016; Fan, Lu and Luo, 2023), railroads (Donaldson and Hornbeck, 2016; Donaldson, 2018), and ports (Ducruet, Juhász, Nagy and Steinwender, 2024).

Section 2 derives the structural gravity equation with endogenous communication; Sections 3–5 present the data, the identification strategy, and the estimates, including the lawsuit-based evidence on the mechanism; Section 6 closes the model in general equilibrium and conducts the counterfactual analysis of three episodes: the 2002–2011 improvement in air connectedness, the COVID-19 disruption from 2019 to 2020, and the closure of Russian and Ukrainian airspace. Section 7 concludes.

2 Model

In this section, we build a general equilibrium model of sourcing with endogenous communication efforts. Our model combines key elements from Eaton and Kortum (2002) and Furusawa et al. (2017), delivering a structural gravity equation that connects communication costs to trade flows.

There are multiple countries in the world, denoted by o or d, and we denote the set of countries with $\mathcal{N}$. In each country, there are multiple sectors denoted by s or k, and we denote the set of sectors with $\mathcal{S}$. All sectors feature perfect competition, and producers price at marginal costs. In each destination d and sector s, final goods producers combine intermediate goods from various locations and sell their output to consumers. They combine a unit continuum of intermediate goods to produce their output as follows:

$$Q_s = \left(\int_0^1 x(\omega)^{\frac{\rho_s - 1}{\rho_s}} d\omega \right)^{\frac{\rho_s}{\rho_s - 1}}, \tag{1}$$

where $x(\omega)$ denotes the quantity of intermediate goods variety ω, and ρ_s is the elasticity of substitution among different varieties. It will prove convenient to define $\tilde{\rho}_s \equiv \frac{\rho_s}{\rho_s - 1} \in (1, \infty)$; a lower ρ_s (higher $\tilde{\rho}_s$) means that input varieties are more differentiated. Each variety ω can be sourced from a supplier located in any country o. A final goods producer sources ω from the lowest-cost supplier to whom it has access. An intermediate goods producer in an origin country o draws its productivity $z_{os}(\omega)$ to produce any variety from a Fréchet distribution with a cumulative distribution function $F_{os}(z) = \exp(-T_{os} z^{-\theta_s})$, where T_{os} can be seen as the "productivity" of intermediate goods producers in country o and sector s. The parameter θ_s governs the dispersion of productivity draws. We assume $\theta_s > \rho_s - 1$ to ensure a well-defined price index. The unit cost of an input supplier is then $c_{os} t_{od} / z_{os}(\omega)$, where t_{od} is the conventional iceberg trade cost and c_{os} is the unit cost of the input bundle of sector s, country o. In our empirical analysis, we do not need to specify c_{os} because it is absorbed by origin-sector fixed effects. In the quantitative analysis in Section 6, we specify c_{os} as a function of wages and prices of intermediate inputs.

Different from the Eaton and Kortum (2002) model, we incorporate endogenous communication that affects the probability, α, that an input meets the buyer's standard. When the input fails to meet the standard, it becomes useless to the buyer and drops out of the integration in equation (1). The probability, α, is the buyer's choice variable and can be interpreted as a communication intensity; the buyer can raise α with an increased communication cost of $e^{m_{od}\alpha}$, which scales the unit cost of the input in a multiplicative way. We refer to m_{od} as the exogenous communication cost. It is specific to an origin-destination country pair but constant among supplier-buyer pairs, conditional on the origin and destination. When it comes to empirical estimation, we parameterize m_{od} as a function of observables. We adopt a simple exponential function for the communication costs to obtain closed-form solutions. In Online Appendix A.2, we discuss more general functional forms for the communication costs and show the robustness of our main predictions.

It will be clear that buyers in country d choose the same communication intensity across all sellers in country o, and that they source inputs from the suppliers offering the lowest effective price, with the risk of product failure taken into account. To understand this, let us consider the buyer's optimal communication and sourcing strategies in the hypothetical case that the buyer in d sources all input varieties $\omega \in [0,1]$ from a sole country o. The buyer may well want to vary its communication intensity across sellers, depending on their productivity through the prices that they offer, as well as their location. Thus, we allow the buyer to choose α as a function of an input price p, i.e., $\alpha = \alpha_{ods}(p)$. Then, it follows from the convention that, by the law of large numbers, the probability $\alpha_{ods}(p)$ is also the fraction of inputs that arrive intact among the inputs sourced from the sellers with their product price of p, and that the probability distribution for the price of an input coincides with the distribution of prices for those intact inputs. Therefore, for any given $\alpha_{ods}(\cdot)$, we can write the cost minimization problem that faces the buyer as

$$\begin{aligned} \min_{\{x_{ods}(\cdot)\}} \quad & \int_0^\infty p x_{ods}(p) dG_{ods}(p) \\ \text{s.t.} \quad & \left[\int_0^\infty \alpha_{ods}(p) x_{ods}(p)^{\frac{\rho_s-1}{\rho_s}} dG_{ods}(p)\right]^{\frac{\rho_s}{\rho_s-1}} \geq 1, \end{aligned} \tag{2}$$

where $G_{ods}(\cdot)$ is the probability distribution of p induced by the idiosyncratic productivity shock $z_{os}(\omega)$. We solve this problem to obtain input demands as

$$x_{ods}(p) = p^{-\rho_s} \alpha_{ods}(p)^{\rho_s} \left[\int_0^\infty \left(\alpha_{ods}(p)^{-\tilde{\rho}_s} p\right)^{1-\rho_s} dG_{ods}(p)\right]^{-\tilde{\rho}_s},$$

which entails the price index of the composite input sourced from country o as

$$P_{ods} = \left[\int_0^\infty \left(\alpha_{ods}(p)^{-\tilde{\rho}_s} p \right)^{1-\rho_s} dG_{ods}(p) \right]^{\frac{1}{1-\rho_s}}. \tag{3}$$

Note that the risk of product failure effectively raises the price by a multiple of $\alpha_{ods}(p)^{-\tilde{\rho}_s}$.

But this effective price increase is endogenous. We now consider a communication intensity that is not only chosen optimally but can also potentially differ across the varieties, with different prices that they offer, conditional on being sourced from country o, which we represent by $\alpha_{ods}(p(\omega))$. Since the price equals the marginal cost of input production, the buyer minimizes the effective price, $\alpha_{ods}(p(\omega))^{-\tilde{\rho}_s} p(\omega)$, variety by variety:

$$\min_{\alpha_{ods}(p(\omega))} \alpha_{ods}(p(\omega))^{-\tilde{\rho}_s} e^{m_{od}\alpha_{ods}(p(\omega))} c_{os} t_{od} / z_{os}(\omega), \tag{4}$$

where $e^{m_{od}\alpha_{ods}(p(\omega))} c_{os} t_{od} / z_{os}(\omega)$ is the supplier ω's production cost, as well as its factory-gate price due to perfect competition. Solving this problem, the buyer chooses the optimal communication intensity as

$$\alpha_{ods}(p(\omega)) = \frac{\tilde{\rho}_s}{m_{od}},$$

where $\tilde{\rho}_s$ represents the degree of product differentiation.[4] It is readily seen that the optimal α_{ods} does not depend on ω; the communication intensities are the same across all sellers within the same country o. Now, we can substitute the optimal $\alpha_{ods}(p(\omega))$, which we simply call α_{ods}, into the expression for the effective price to obtain

$$\alpha_{ods}^{-\tilde{\rho}_s} p = \left(\frac{e m_{od}}{\tilde{\rho}_s} \right)^{\tilde{\rho}_s} c_{os} t_{od} / z_{os}(\omega). \tag{5}$$

We see that $(e m_{od}/\tilde{\rho}_s)^{\tilde{\rho}_s}$ is the extra "iceberg" cost due to communication and quality control. At the optimum, input failures raise costs by $\alpha_{ods}^{-\tilde{\rho}_s} = (m_{od}/\tilde{\rho}_s)^{\tilde{\rho}_s}$, and achieving this reliability requires a communication outlay of $e^{m_{od}\alpha_{ods}} = e^{\tilde{\rho}_s}$.

It is worth being explicit about who pays for the communication. The communication and quality-control activity is carried out by the supplier with its own production factors: in the price $e^{m_{od}\alpha_{ods}} c_{os} t_{od} / z_{os}(\omega)$, the communication factor multiplies the supplier's unit cost, so delivering one unit at reliability α_{ods} requires $e^{m_{od}\alpha_{ods}}$ times the resources that production alone would require. Because intermediate goods producers are perfectly competitive and price at marginal cost, the buyer compensates the supplier for these extra resources through

[4] We assume that the level of m_{od} is sufficiently high so that $\tilde{\rho}_s / m_{od}$ is below one and we do not have a corner solution for the optimal $\alpha_{ods}(p(\omega))$.

the price she pays. A higher success probability thus requires the supplier to use more resources in production. Meanwhile, a lower exogenous communication cost m_{od} reduces the extra resources needed to attain the same α_{ods}: direct flights make it easier for the supplier to understand the buyer's needs and thereby lower the cost of delivering products of the same reliability.

Having derived the buyer's communication and sourcing strategies when they source all the inputs from one country, we can derive their general strategies for sourcing inputs worldwide. As in the original EK model, buyers source each input ω from the location that offers them the lowest effective price, i.e., they solve $\min_o \left\{(em_{od}/\tilde{\rho}_s)^{\tilde{\rho}_s} c_{os} t_{od}/z_{os}(\omega)\right\}$ for each ω. Now, under the restriction that $\theta_s > \rho_s - 1$, we can apply the properties of the Fréchet distribution and obtain the unit cost for the composite input bundle in sector s in country d as

$$P_{ds} = \gamma_s \left[\sum_o T_{os} \left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od}\right)^{-\theta_s}\right]^{-1/\theta_s}, \tag{6}$$

where $\gamma_s \equiv \left(\frac{e}{\tilde{\rho}_s}\right)^{\tilde{\rho}_s} \left[\Gamma\left(\frac{\theta_s+1-\rho_s}{\theta_s}\right)\right]^{1/(1-\rho_s)}$ is a constant and Γ denotes the Gamma function.

It is immediate that the parameter ρ_s plays an important role in our model: a higher communication cost m_{od} is more costly for inputs with lower ρ_s (i.e., higher $\tilde{\rho}_s$). This is in contrast with Eaton and Kortum (2002), where the elasticity of substitution only affects the price levels in all locations through the constant γ_s, thus can be ignored in their analysis as long as $\theta_s > \rho_s - 1$.

The trade flow from o to d can then be written as

$$X_{ods} = \frac{T_{os} \left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od}\right)^{-\theta_s}}{\sum_o T_{os} \left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od}\right)^{-\theta_s}} X_{ds},$$

where X_{ds} is total absorption of destination d, sector s. Taking the log of both sides, we obtain the following gravity equation

$$\log X_{ods} = \delta_{ds} + \log T_{os} - \theta_s \log c_{os} - \theta_s \tilde{\rho}_s \log m_{od} - \theta_s \log t_{od}, \tag{7}$$

where δ_{ds} is the destination-sector fixed effect. A key feature of our model is that communication costs, m_{od}, affect trade flows differently from conventional trade costs, t_{od}. In particular, the elasticity of trade flows with respect to t_{od} is θ_s, the EK parameter that governs the inverse dispersion of productivity draws. In contrast, the elasticity of trade flows with respect to m_{od} is $\theta_s \tilde{\rho}_s$, which also depends on the elasticity of substitution between different varieties of inputs. This is crucial for isolating the role of communication costs. In

Sections 4 and 5, we use a regression specification based on equation (7) and estimate the impact of direct flights on m_{od}. As another advantage of our model, we can apply the "exact hat algebra" in Dekle et al. (2007) to perform counterfactual analysis in general equilibrium. We analyze the equilibrium impact of direct flights and their welfare implications in Section 6.

Our model shares a common mathematical structure with Fieler and Eaton (2025): in both, a shifter in demand raises the effective quantity of a sourced variety, and raising this shifter scales the variety's unit cost. A natural question is whether reinterpreting α as a quality index, in the same spirit as Fieler and Eaton (2025), would deliver the same gravity equation. We examine this in two steps in the Online Appendix. In Online Appendix A.1, we show that the way α enters demand determines the elasticity of trade with respect to communication. In our model, α is a probability and enters the integral (2) linearly. Letting α enter as the "substitutable" quality as in Fieler and Eaton (2025), with a power of $(\rho_s - 1)/\rho_s$, would lead to a communication elasticity that does not vary with ρ_s, while letting α enter as a demand shifter with a power of $1/\rho_s$, as in Broda and Weinstein (2006), would lead to a different elasticity but maintain the property that the elasticity decreases with ρ_s. The difference between the first two cases is what a failure destroys: in our model a variety that fails is lost entirely, whereas substitutable quality corresponds to a fraction of the units of the variety failing. Our empirical evidence below is consistent with the success probability interpretation. In Online Appendix A.2, we consider more general forms in which α affects the unit costs. We show that the communication elasticity decreases with ρ_s as long as the unit cost is log-supermodular in m_{od} and α, and our main specification $e^{m_{od}\alpha}$ is a special case that delivers closed-form solutions.[5]

3 Data

In this section, we describe the datasets and construct the key variables in our empirical and quantitative analysis.

3.1 Direct Flight Data

Our data on direct flights come from the International Civil Aviation Organization (ICAO). More specifically, we use the Traffic by Flight Stage (TFS) module, which provides informa-

[5] In Fieler and Eaton (2025), the production of quality contains no bilateral, origin–destination term, so bilateral frictions enter their gravity equation only through a conventional iceberg trade cost—the analog of the term $\theta_s \log t_{od}$ in (7). The comparison in the Online Appendix instead posits that communication costs directly affect the production of quality in their framework, and asks whether this delivers the same gravity equation as ours.

tion about individual flight stages of international scheduled services over several decades[6]. It covers the name of the city and country where the non-stop flight takes off and lands and also includes information such as the number of flights operated and average passenger capacity. To exclude the direct impact of international flights on air freight, we only use data on flights with a positive number of seats available.

We focus on regular round-trip flights between cities and do not consider as connected city pairs with only a few direct flights a year or those with direct flights in one direction. Following Campante and Yanagizawa-Drott (2018), we define a city pair as "connected" if there are weekly round-trip flights between these two cities (at least 52 direct flights in both directions). The TFS data cover around 2,000 cities, and we find around 7,000 city pairs with weekly direct flights.

3.2 Bilateral Trade Data

We obtain bilateral trade data from BACI (Guillaume and Zignago, 2010). BACI provides data on bilateral trade flows for more than 200 countries at the HS 6-digit product level. We use 2016 as the main year of analysis. To focus our analysis on the largest importing countries and make our sample size more manageable, we select the top 100 countries in total imports, which account for more than 95% of world trade.

3.3 Business Lawsuits

Our model implies that the probability of product failure increases with communication costs. If direct flights reduce communication costs, we expect they reduce the probability of product failure. Since it is difficult to obtain measures of product failure between buyers and suppliers, we use the occurrence of business lawsuits as a proxy.

We obtain data from the Lexis+ Hong Kong database provided by LexisNexis, and follow similar procedures described in Boehm (2022) to extract cases involving the keyword "defective product", which connects closely with the notion of product failure in our model.[7] We focus on cases that occurred in the federal district courts in the thirteen judicial circuits of the United States between 2011 and 2017. We download the full case documents and, based on the parties listed in each document's header, keep only cases in which firms appear

[6] See https://data.icao.int/newDataPlus/Dataplus/App_TrafficFlightStage.

[7] Boehm (2022) extract cases on general business disputes using the keyword "breach of contract". We have also examined this keyword, but many of the resulting cases involve services or disputes unrelated to manufactured products, so the sample maps less well into our model. We therefore report the corresponding results only as a robustness check in Online Appendix B.7. The two samples overlap, but neither nests the other.

on both sides; when a case involves multiple plaintiffs or defendants, we count it by the share of plaintiff–defendant pairs that are firm pairs.[8] To obtain the locations of the firms involved in these cases, we use the names of plaintiffs and defendants and match them to firms in ORBIS, a global firm-level database maintained by Moody's. Note that we do not use the country or city of the firms directly involved in these cases, but rather the location information of these firms' global ultimate owners (GUO). There are two reasons for this choice. First, the firm names in the Lexis database are short, and they can be matched to multiple subsidiaries of a single GUO. It is hard to determine which subsidiary is the actual plaintiff or defendant. Second, when a US firm is involved in a lawsuit with a US subsidiary of a German multinational firm, for example, we see the business dispute affected by the communication costs between the US firm and the German headquarter. Finally, we allocate each GUO to a $1^\circ \times 1^\circ$ grid by geocoding its city. Online Appendix B.3 details the construction of the firm-vs-firm sample, the matching procedure, and the weighting scheme that defines the effective number of cases between grids.

Excluding cases between firms from the same country, we end up with a total of 559 international defective-product lawsuits between firms, distributed across 437 country pairs. Since the lawsuits are filed in the US, it is not surprising that either the plaintiff or the defendant is often from the US. However, we also observe cases between firms from other countries, especially those with large economies, such as Germany and the United Kingdom: in 339 country pairs, both the plaintiffs and the defendants are located outside the US. We provide the list of the top 30 country pairs with the most litigation cases in Online Appendix Table B-1. Our empirical results, however, are robust when we exclude the US from the sample.

3.4 Firm Ownership Links

Direct flights may increase the number of business lawsuits between two locations because they reduce communication costs and increase trade or other business activities, even if they reduce the *probability* of product failure. Therefore, we also need to examine the impact of direct flights on total business activities between two locations. At the country-pair level, we can use bilateral trade data for this purpose. However, at the grid-pair level, we do not have trade data, and we use the number of firm ownership links between two grids as a proxy for business activities. We identify an ownership link from grid j to grid i if the GUO is located in grid j and it has a subsidiary in grid i as recorded in ORBIS.

We obtain a total of 270,096 foreign ownership links between grids in different countries

[8] Online Appendix B.3 describes the sample construction in detail and presents a textual and case analysis of the case documents.

from 2011 to 2017. These links are distributed across 2,654 country pairs and 59,932 grid pairs. We provide the list of the top country pairs with the most foreign ownership links in Online Appendix Table B-2.

3.5 Additional/Auxiliary Variables

Our structural gravity equation (7) involves the usual trade elasticity θ_s and the elasticity of substitution ρ_s. When implementing the estimation, we use existing estimates of these parameters from the literature. For each product (sector) s, we calculate $\tilde{\rho}_s = \frac{\rho_s}{\rho_s - 1}$ using the elasticity of substitution estimated by Soderbery (2015). The original estimates of ρ_s are at the HS 8-digit level, and we convert them to HS 6-digit levels using simple averages. For the Fréchet parameter θ_s, we use the estimates in Caliendo and Parro (2015) in our main specification and those in Giri, Yi and Yilmazkuday (2020) for robustness checks.[9] Both papers estimate θ_s at more aggregated industry levels. When mapping their estimates to our sample, we assume that all products that belong to the same industry have the same θ_s.

4 Empirical Specifications

This section describes our empirical methods. The endogeneity issue arises because the existence of direct flights may be affected by the intensity of commercial activities between two locations. To address this issue, we construct an instrumental variable to estimate the structural gravity equation and obtain the causal effect of direct flights on communication costs.

4.1 Country-pair Level Analysis on Trade Flows

4.1.1 Connection Index between Countries

Since trade data are more geographically aggregated than flight data, we refer to the method of calculating the effective distance between country pairs (Mayer and Zignago, 2011) to construct a country-to-country connection index. In particular, we divide each country into $1^\circ \times 1^\circ$ grids, approximately 69 miles $\times$ 69 miles cells.[10] We obtain the population of each grid

[9] Caliendo and Parro (2015) estimate trade elasticities using trade flows and tariffs for 20 tradeable goods sectors (approximately two-digit ISIC industries). In contrast, Giri et al. (2020) use the micro price data from Eurostat surveys to construct the sector prices inclusive of trade costs to estimate trade elasticities for ISIC three-digit manufacturing industries. Their approach is based on the price implications in Eaton and Kortum (2002) and follows the SMM method developed by Simonovska and Waugh (2014) that addresses small sample biases. Both approaches are consistent with our sectoral Eaton-Kortum model.

[10] In practice, the area of each grid depends on its location. One degree of latitude is equivalent to approximately 69 miles anywhere on the Earth. At the equator, one longitude is also equivalent to 69 miles.

from the Geographically based Economic (G-Econ 4.0) data.[11] We drop grids not covered by G-Econ and assume that there are no economic activities in those locations.

We next define connectedness between any two grids. We utilize the Google Map Geocoding API to obtain the longitude and latitude of cities in the TFS data and assign each city to a $1^\circ \times 1^\circ$ grid based on its coordinates.[12] We calculate the connection index between countries o and d, κ_{od}, using the following formula:

$$\kappa_{od} = \sum_{j \in o} \omega_j \sum_{i \in d} \omega_i \kappa_{ji}, \tag{8}$$

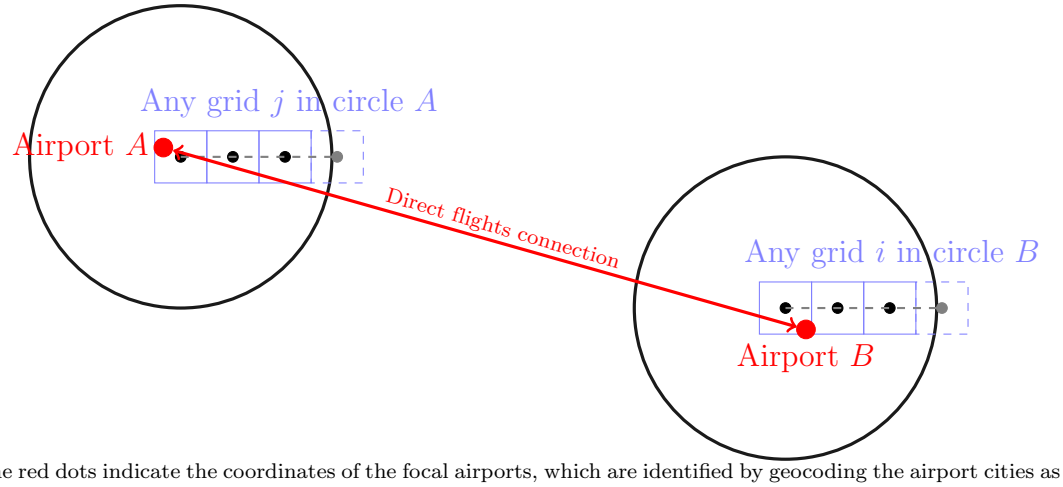
where we abuse the notation slightly and use o and d also to denote the set of grids in the two countries. j denotes a grid in the origin country, and i denotes a grid in the destination country. κ_{ji} is a dummy variable, which equals one if the grid centers of j and i are within a 200-mile radius of two TFS cities that are connected by weekly direct flights. The weights, ω_j and ω_i, are the share of population of grid j in country o and that of grid i in country d, respectively. We use population data from 1990, the earliest year for which G-Econ reports population data, to calculate the weights. Figure 1 illustrates how we assign values to κ_{ji}. We draw a circle of 200 miles around the center of the grid where Airport A is located (Circle A) and a circle of 200 miles around the center of the grid where Airport B is located (Circle B). As long as Airports A and B are connected by direct flights, any grid j whose center is inside Circle A is considered to be connected to any grid i whose center is inside Circle B, i.e., $\kappa_{ji} = \kappa_{ij} = 1$. We highlight these grids with black centers and solid edges in Figure 1.

But the distance of one longitude gradually shortens when moving from the equator to the poles.

[11] https://gecon.yale.edu/data-and-documentation-g-econ-project.

[12] Since the Google Maps API may return multiple pairs of coordinates for a city, we obtain the unique coordinates of each TFS airport city using two additional steps. First, we use each country's shape file and remove coordinates that fall outside the corresponding country of each city. Second, for cities with multiple coordinates, if the maximum distance between these coordinates is less than 100 miles, we use the simple average of these coordinates. If the maximum distance is above 100 miles, we manually search for the coordinate that contains the largest airport in the city.

Figure 1: Illustration of Connection between Grid j and Grid i



Notes. The red dots indicate the coordinates of the focal airports, which are identified by geocoding the airport cities as reported in the TFS data. The squares represent $1° \times 1°$ grids, and the black/grey dots are the grid centers. We draw a circle of 200 miles radius around the grid centers that are closest to the airport locations. The grid pairs whose centers fall within the circles (with black centers and solid edges) are considered connected if the two airports have established a direct flight connection.

4.1.2 Estimation Equation

We use the structural gravity equation (7) to estimate the effect of direct flights on communication costs, m_{od}. We parameterize the communication costs as a function of the connection index in equation (8):

$$m_{od} = e^{\beta_0 - \beta \kappa_{od}}.$$

The constant β_0 scales the level of the communication costs and ensures that m_{od} is high enough for the optimal communication intensity to be an interior solution (footnote 4). It is absorbed by the fixed effects in the estimation below and cancels from the changes in communication costs in the counterfactual analysis, so we do not need to estimate it. In addition, we parameterize the trade costs, t_{od}, using the fraction of grid pairs within a certain distance range. The distance between any two grids falls into one of the seven groups: 0 - 1K miles, 1K - 2K miles, 2K - 4K miles, 4K - 5.5K miles, 5.5K to 6.5K miles, 6.5K - 8K miles and above 8K miles. We calculate the population-weighted share of grid pairs that fall in, for example, 1K to 2K miles, as follows

$$\lambda_{od,1} \equiv \sum_{j \in o} \omega_j \sum_{i \in d} \omega_i \mathbb{1}\left(d_{ij} \in (1K, 2K]\right).$$

We construct the other shares similarly and denote them as $\lambda_{od,2}, \ldots, \lambda_{od,6}$. We do not need the share of grid pairs in the range of 0 - 1K miles because it will be perfectly collinear with the other six shares. We assume that the log of trade costs can be written as $\log t_{od} \equiv$

$\sum_{n=1}^{6} \gamma_n \lambda_{od,n}$. We prefer this parameterization of trade costs based on grid-pair-level distance because it aligns well with our approach to measuring air connectedness and captures rich topological features of the origin and destination countries.

Substituting the expressions of m_{od} and t_{od} in the gravity equation (7), we obtain our estimation equation

$$\log X_{ods} = \beta \kappa_{od} \times \theta_s \tilde{\rho}_s - \sum_{n=1}^{6} \gamma_n \lambda_{od,n} \times \theta_s + \delta_{ds} + \delta_{os} + \epsilon_{ods}, \tag{9}$$

where δ_{ds} and δ_{os} are destination-sector and origin-sector fixed effects, and ϵ_{ods} is the error term. We estimate equation (9) on the 2016 cross-section: the trade flows and the connection index both refer to 2016.[13]

4.1.3 Identification using Instrumental Variables

The simplest approach to estimate equation (9) is to apply Ordinary Least Squares (OLS). However, such a method suffers from an endogeneity problem. For example, it is possible that unobservable factors in the error term ϵ_{ods} may increase the trade volume and also cause the two countries to have more direct flights (higher κ_{od}). This violates the orthogonality condition $\mathbb{E}(\epsilon_{ods} | \kappa_{od} \times \theta_s \tilde{\rho}_s, \text{controls}) = 0$. In this section, we introduce an instrumental variable (IV) approach following Campante and Yanagizawa-Drott (2018). The key to the IV is a discontinuity of air connectedness between cities at 6,000 miles due to regulations.

As discussed in Campante and Yanagizawa-Drott (2018), the United States and Europe restrict the hours of operation for flight crew members within a 24-hour period, resulting in additional crew on any flight exceeding 12 hours, which corresponds to approximately 6,000 miles of flight distance. Combined with the arrival of new aircraft that could afford such long-haul flights in 1989, the later years saw a significant increase in the number of flights, primarily within the range of 4,600 to 6,000 miles, but not in the range above 6,000 miles. Therefore, there is a discontinuous drop in direct flights at around 6,000 miles. In Online Appendix B.1, we replicate their findings using the 2016 TFS data.

We then construct an instrumental variable (IV) for the country-to-country connection index. Our IV has the same spirit as the IV for city-level connectedness in Campante and Yanagizawa-Drott (2018). In particular, for two countries, o and d, we focus on the grid

[13] The TFS schedules allow us to construct κ_{od} in other years as well. The gravity estimation uses only the 2016 cross-section. The analysis of business lawsuits and foreign ownership links uses the average connectedness between 2011 and 2017, matching the time span of those outcomes (Section 4.2). The over-time variation in connectedness is not used in the estimation; it appears only in Section 6, where we feed observed changes in connectedness into the counterfactual analysis.

pairs with a distance in the range of 5,500 to 6,500 miles. Within this set of grid pairs, we compute the population-weighted share of pairs that are below 6,000 miles. In particular, our IV is

$$z_{od} = \sum_{j \in o, i \in d} \omega_{ji} \mathbb{1}\left(d_{ji} \leq 6K | d_{ji} \in (5.5K, 6.5K]\right) \tag{10}$$

where d_{ji} is the distance between grid j in country o and grid i in country d, and the population weights are calculated as $\omega_{ji} \equiv \frac{\text{pop}_j \times \text{pop}_i}{\sum_{j \in o, i \in d} \text{pop}_j \times \text{pop}_i \times \mathbb{1}(d_{ji} \in (5.5K, 6.5K])}$.[14] We assume that conditional on being in the range of 5,500 to 6,500 miles, the share of grid pairs that fall below 6,000 miles is as good as randomly assigned.

Since the IV is a conditional probability, it is only defined for country pairs with at least one pair of grids between 5,500 and 6,500 miles. This restriction reduces the total number of country pairs in our sample from over 10,000 to around 4,000. It is desired because we want to exclude country pairs that are either too far apart or too close to each other. For these country pairs, their connection index κ_{od} is determined by the overall geographical distance, not by the regulation that causes the discontinuity of direct flights at 6,000 miles.

In our origin-destination-product level regressions, the endogenous regressor is $\kappa_{od} \times \theta_s \tilde{\rho}_s$ instead of κ_{od}. Therefore, we multiply z_{od} by $\theta_s \tilde{\rho}_s$ and use it as the IV in the gravity regression. Formally, the exclusion restriction for the IV is

$$\mathbb{E}\left(\epsilon_{ods} | z_{od} \times \theta_s \tilde{\rho}_s, \text{controls}\right) = 0.$$

Note that z_{od} is time-invariant by construction: it depends only on geography and predetermined population weights. The identifying variation is therefore cross-sectional—conditional on the fixed effects and the distance controls, β is identified by comparing trade flows across country pairs whose 2016 connectedness differs with z_{od}.

We provide summary statistics of κ_{od} and z_{od} across country pairs in Table 1. Across all country pairs, the average population-weighted air connectedness is 0.158. Among the country pairs with at least one pair of grids in the 5,500- and 6,500-mile range, the average connection index is lower (0.065). In general, the connection index contains many zeros, and 20% of the country pairs are connected by weekly direct flights. The instrumental variable, z_{od}, has an average of 0.544. It means that within the grid pairs that are between 5,500 and 6,500 miles, approximately half of the grid pairs have a distance of less than 6,000 miles. This finding is consistent with our assumption that the share of grid pairs falling below 6,000 miles is as good as randomly assigned, given that they are within the 5,500- and 6,500-mile

[14] The weights here are slightly different from those in equation (8). When defining the connection index, we do not restrict the grid pairs to be between 5,500 and 6,500 miles and calculate the weights as $\omega_j \times \omega_i = \frac{\text{pop}_j \times \text{pop}_i}{\sum_{j \in o, i \in d} \text{pop}_j \times \text{pop}_i}$.

range.

Table 1: Summary Statistics

	Obs.	mean	Std. Dev.	min	p1	p10	p25	p50	p75	p90	p99	max
κ_{od} (Full Sample)	9900	0.158	0.298	0	0.000	0.000	0.000	0.000	0.174	0.731	1.000	1.000
κ_{od} (Sample in (5.5K,6.5K))	3024	0.065	0.162	0	0.000	0.000	0.000	0.000	0.000	0.261	0.764	0.997
z_{od}	3024	0.544	0.424	0	0.000	0.000	0.025	0.635	1.000	1.000	1.000	1.000
# Grid Pairs	3024	66064.7	322668.3	6	48	340	1344.5	5929	29036.5	122101	904428	7752241
# Grid Pairs in (5.5K,6.5K)	3024	11761.4	52447.3	1	2	40	210	1142.5	5901.5	19779	184183	978547

Notes. The first row of the table reports summary statistics of the connection index κ_{od} across all country pairs. The remainder of the table reports κ_{od}, the instrument z_{od}, the number of all grid pairs and of grid pairs in the 5,500- to 6,500-mile range in a subsample of country pairs with at least one grid pair in the 5,500- to 6,500-mile range.

Table 2 lists the top 20 country pairs with the highest connectedness between 5,500 and 6,500 miles and the country pairs close to the average connectedness. We can see that economies with concentrated populations or high income are more likely to have high connectedness. For instance, Hong Kong and Singapore have high connection indices with many European countries. The connection index between Turkey and Indonesia is 0.066, and that between the USA and Russia is 0.064, closest to the average of country pairs in our sample.

Table 2: List of the Most Connected Country Pairs, 2016

Rank	Country 1	Country 2	κ_{od}	# Grid Pairs	# Grid Pairs in (5.5K,6.5K)	z_{od}	# Grid Pairs in (5.5K,6.0K)
1	Singapore	Denmark	0.997	91	72	0.000	0
2	Singapore	Switzerland	0.987	48	37	0.000	0
3	Hong Kong	Switzerland	0.987	48	48	1.000	47
4	Hong Kong	United Kingdom	0.968	258	228	0.055	90
5	Singapore	Netherlands	0.961	57	19	0.000	0
6	Hong Kong	Netherlands	0.961	57	45	1.000	45
7	Singapore	Germany	0.910	213	198	0.000	0
8	Hong Kong	Germany	0.910	213	182	1.000	182
9	Singapore	Finland	0.865	306	306	1.000	298
10	Hong Kong	Australia	0.865	2454	3	1.000	3
11	Sri Lanka	United Kingdom	0.858	1118	740	1.000	740
12	Costa Rica	United Kingdom	0.858	946	79	1.000	79
13	Netherlands	Panama	0.807	399	288	1.000	288
14	Qatar	Japan	0.783	721	22	1.000	22
15	Finland	Japan	0.772	12241	137	1.000	137
16	Hong Kong	Austria	0.764	81	61	1.000	61
17	United Arab Emirates	Australia	0.761	17178	8021	0.928	4286
18	Hong Kong	Italy	0.757	255	255	0.954	226
19	Singapore	Italy	0.757	255	241	0.036	20
20	Netherlands	Japan	0.753	2281	1512	0.987	1251
286	Turkey	Indonesia	0.066	44031	19303	0.601	9722
287	USA	Russia	0.064	4857373	978547	0.552	638955

Notes. this table reports the most connected country pairs and the country pairs with average connectedness. We rank country pairs by the connectedness index, κ_{od}. Note that this index and our instrument variable z_{od} are symmetric, i.e., $\kappa_{od} = \kappa_{do}$ and $z_{od} = z_{do}$. We therefore only keep one country pair in the table, and the country with a smaller area comes as "Country 1".

In Table 3, we show that our IV positively predicts the connection index κ_{od}. In Column (1), we regress κ_{od} on z_{od} without controlling for any other variables. We see a coefficient of 0.065 with a standard error of 0.007 and an R-squared of 0.029. For many country pairs, the overall air connectedness κ_{od} is also affected by grid pairs that are much less likely to be connected (more than 6,500 miles apart) and/or those that are much more likely to be connected (less than 5,500 miles apart). Nevertheless, the conditional probability z_{od} has a statistically significant effect on overall connectedness. The effect is smaller but still significant when we include origin and destination fixed effects in Column (2). Since not all country pairs in our IV sample are among the top 100 importers in 2016, we run the same regressions using the pairs that are also in our trade sample in Columns (3) and (4). The results are robust. Note that these regressions are not the actual first stages of the IV regressions of equation (9). We use the product of $\theta_s\tilde{\rho}_s$ and z_{od} as an IV for $\theta_s\tilde{\rho}_s \times \kappa_{od}$.

Table 3: Predicting overall air connectedness using IV

Dep. Var: Connection Index$_{od}$	IV Sample		IV & Trade Sample	
	(1)	(2)	(3)	(4)
$\sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K] \text{ miles}\right)$	0.065^a	0.044^a	0.065^a	0.045^a
	(0.007)	(0.007)	(0.007)	(0.007)
Origin FE	N	Y	N	Y
Dest. FE	N	Y	N	Y
Observations	3024	3024	2959	2959
# of Destination-Origin groups	3024	3024	2959	2959
F-stat	84.64	41.75	82.86	41.57
R-squared	0.029	0.453	0.029	0.454

Notes. The IV sample refers to the country pairs with at least some grid pairs that are 5,500 to 6,500 miles apart. The IV & trade sample requires the country pairs to appear in the 100 countries selected in our trade data and to have positive trade flows in 2016. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10.

4.2 Grid-pair Level Analysis on International Business Lawsuits and Ownership Links

In this section, we turn to the specifications used to estimate the impact of direct flights on business lawsuits, taking into account their effect on business activities at the grid-pair level. An advantage of the grid-pair level analysis is that we can use the discontinuity of direct flights at 6,000 miles to perform a regression discontinuity (RD) design.

We first employ a sharp RD design to examine the "reduced-form" relationship between distance and the number of business lawsuits and activities, with the latter measured by the number of foreign ownership links. We use the following regression specification:

$$Y_{ji} = \beta_0 + \beta_1 \mathbb{1}\left(d_{ji} \leq 6K\right) + \beta_2 (d_{ji} - 6K) + \beta_3 (d_{ji} - 6K) \times \mathbb{1}\left(d_{ji} \leq 6K\right) + \epsilon_{ji}, \tag{11}$$

where we use Y_{ji} to denote the number of business lawsuits or the number of foreign ownership links between grid j and grid i, and d_{ji} is the distance between the two grids.[15] The first term, $\mathbb{1}\left(d_{ji} \leq 6K\right)$, is an indicator variable that equals one if the distance between the two grids is less than or equal to 6,000 miles. The second and third terms allow for linear relationships between the outcome variables and distance on either side of the 6,000-mile threshold, potentially with different slopes. We limit our analysis to grid pairs with distances within a small bandwidth of the threshold. We provide robustness tests using second-order polynomials, estimates with different sample bandwidths, and different kernel function choices. We adopt robust standard errors as recommended by Imbens and Lemieux (2008).

[15]The number of business lawsuits is the effective (weighted) number of defective-product cases; Online Appendix B.3 describes the matching to the firms' global ultimate owners and the weighting scheme.

As in Campante and Yanagizawa-Drott (2018), we can test whether there is a first-stage relationship between the indicator of distance being less than 6,000 miles and whether two grids are connected by direct flights, simply by using the grid-pair level connectedness index $\bar{\kappa}_{ji}$ as the dependent variable. To be consistent with the time span of business lawsuits and foreign ownership links, we use the average connectedness $\bar{\kappa}_{ji}$ between 2011 and 2017. Under the exclusion restriction that the distance around the 6,000-mile threshold affects the number of business lawsuits or foreign ownership links only through direct flight connectedness, we can use $\mathbb{1}(d_{ji} \leq 6K)$ as an instrumental variable for $\bar{\kappa}_{ji}$ and apply the "fuzzy" RD method to obtain the impact of air connectedness, $\bar{\kappa}_{ji}$, on the outcome variables, Y_{ji}.

5 Empirical Results

In this section, we present our main empirical results. We first investigate the impact of air connectedness on trade flows by estimating the structural gravity equation (9) using OLS and IV methods. We then turn to the grid-pair level analysis and examine the impact of direct flights on business lawsuits and foreign ownership links using a regression discontinuity design.

5.1 Impact of Air Connectedness on Trade

Table 4 reports our baseline regression results: the impact of the connection index, κ_{od}, on the sector-level trade volume. Column (1) reports an OLS regression (9) using the full sample, including all origin-destination-product level trade flows among the top 100 trading economies. Consistent with our model, we find a positive coefficient $\hat{\beta} = 0.015$ with a standard error of 0.002. This suggests that air connectedness between two countries reduces communication costs and promotes trade for products with a lower dispersion of productivity draws (higher θ_s) or a lower elasticity of substitution (lower ρ_s thus higher $\tilde{\rho}_s$). The coefficients of the control variables $\theta_s \times \lambda_{od,n}, 1 \leq n \leq 6$ capture the impact of distance on trade costs and trade flows. As discussed earlier, the instrumental variable z_{od} is only defined for a subset of country pairs with at least some grid pairs with a distance between 5,500 and 6,500 miles. We perform the OLS regression for this subsample in Column (2) and obtain a slightly larger effect of air connectedness.

We present our preferred specification, the IV regression, in Column (3). We instrument the endogenous variable $\theta_s\tilde{\rho}_s \times \kappa_{od}$ with $\theta_s\tilde{\rho}_s \times z_{od}$. The IV estimate of β is 0.026, similar to the OLS estimate in Column (2), with a standard error of 0.012. We report the corresponding first-stage regression in Column (4). Our instrument variable significantly predicts the endogenous variable, and the first-stage Kleibergen-Paap F-statistic is 36.29. According to

the IV estimates, with a ten percentage point increase in air connectedness between the two countries, the trade flow of the product with a median value of $\theta_s\tilde{\rho}_s$ will increase by 2.5%. Products at the 75th percentile will see a 4.1% increase in trade when air connectedness increases by ten percentage points.[16]

The similarity between the IV and OLS estimates is not obvious ex ante, because two sources of bias work in opposite directions. On the one hand, reverse causality and omitted bilateral commercial ties push the OLS estimate upward: airlines are more likely to offer direct flights where demand for business travel is high, and this demand is plausibly strongest for the communication-intensive products from which our coefficient is identified. On the other hand, the connection index is a constructed proxy for the true cost of communication, so measurement error attenuates the OLS estimate toward zero. The evidence below (Table 5) suggests that the upward bias is present but modest in our sample: historical trade flows predict the endogenous variable conditional on our controls, and controlling for them lowers the OLS estimate while leaving the IV estimate essentially unchanged. The remaining upward bias is thus small and roughly offset by attenuation, consistent with the IV point estimate lying slightly above the OLS estimate. Moreover, the IV estimate captures a local average effect for country pairs whose connectedness responds to the discontinuity at around 6,000 miles, so we do not expect it to be exactly equal to the OLS estimate.

[16]The median (75th percentile of) $\theta_s\tilde{\rho}_s$ is 9.51 (15.9) in our sample. The median ρ_s is 2.05.

Table 4: Impact of Air Connectedness on Trade Volume

	Full Sample	Sample in (5.5K,6.5K)		
	OLS	OLS	IV	First Stage
	(1)	(2)	(3)	(4)
Connection Index$_{od} \times \theta_s\tilde{\rho}_s$	0.015^a	0.022^a	0.026^b	
	(0.002)	(0.003)	(0.012)	
$\theta_s\tilde{\rho}_s \times \sum_{ji\in od} w_{ji}\mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$				0.091^a
				(0.015)
$\theta_s \times \sum_{j\in o} w_j \sum_{i\in d} w_i\mathbb{1}\left(d_{ij} \in (1K, 2K]\right)$	-0.223^a	-0.438^a	-0.436^a	-0.580
	(0.008)	(0.093)	(0.094)	(0.394)
$\theta_s \times \sum_{j\in o} w_j \sum_{i\in d} w_i\mathbb{1}\left(d_{ij} \in (2K, 4K]\right)$	-0.317^a	-0.679^a	-0.675^a	-0.790^b
	(0.008)	(0.077)	(0.078)	(0.315)
$\theta_s \times \sum_{j\in o} w_j \sum_{i\in d} w_i\mathbb{1}\left(d_{ij} \in (4K, 5.5K]\right)$	-0.389^a	-0.740^a	-0.735^a	-1.239^a
	(0.008)	(0.077)	(0.079)	(0.308)
$\theta_s \times \sum_{j\in o} w_j \sum_{i\in d} w_i\mathbb{1}\left(d_{ij} \in (5.5K, 6.5K]\right)$	-0.392^a	-0.752^a	-0.747^a	-1.202^a
	(0.008)	(0.077)	(0.080)	(0.312)
$\theta_s \times \sum_{j\in o} w_j \sum_{i\in d} w_i\mathbb{1}\left(d_{ij} \in (6.5K, 8K]\right)$	-0.420^a	-0.761^a	-0.756^a	-1.209^a
	(0.009)	(0.077)	(0.080)	(0.313)
$\theta_s \times \sum_{j\in o} w_j \sum_{i\in d} w_i\mathbb{1}\left(d_{ij} \geq 8K\right)$	-0.442^a	-0.759^a	-0.751^a	-1.671^a
	(0.010)	(0.077)	(0.081)	(0.315)
Dest.-HS6 FE	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y
Observations	5855591	1779937	1779937	1779937
# of Destination-Origin groups	9524	2926	2926	2926
1st-Stage F-stat			36.29	
R-squared	0.563	0.699		0.721

Notes. Each observation is an origin-destination-HS6 product combination. The dependent variable in Columns (1)–(3) is the log of trade flows. The dependent variable in Column (4) is the endogenous variable, $\kappa_{od} \times \theta_s\tilde{\rho}_s$. We obtain estimates of θ_s from Caliendo and Parro (2015). Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10.

Our instrumental variable approach relies on the exclusion restriction that, conditional on the control variables, other factors that influence trade (ϵ_{ods} in equation (9)) are orthogonal to the instrument $z_{od} \times \theta_s\tilde{\rho}_s$. Controlling for historical trade tightens this restriction: it absorbs the persistent components of ϵ_{ods} that may be correlated with the instrument, as well as conventional trade costs.

In Table 5, we re-run the regressions in Table 4, controlling for origin-destination-sector trade flows in 1989 (columns (1)–(3)) or 1992 (columns (4)–(6)). The year 1989 is the earliest year with trade data and the year when new long-haul aircraft were just introduced, so the discontinuous drop in direct flights at around 6,000 miles had not yet emerged (see more evidence in Appendix B.1). Because the earlier trade data contain many more zeros, we transform the control variable with the inverse hyperbolic sine. Historical trade flows positively predict both trade flows and air connectedness in 2016, suggesting an endogeneity problem with OLS regressions (see column (3)); the IV estimates, however, remain similar

to those in our main regressions in both years.

Table 5: Controlling for Historical Trade Flows

	Sample in (5.5K,6.5K)					
	OLS (1)	IV (2)	First Stage (3)	OLS (4)	IV (5)	First Stage (6)
Connection Index $1_{od} \times \theta_s \tilde{\rho}_s$	0.019^a	0.025^b		0.019^a	0.025^b	
	(0.002)	(0.012)		(0.002)	(0.012)	
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$			0.091^a			0.091^a
			(0.015)			(0.015)
Trade in 1989	0.217^a	0.216^a	0.047^b			
	(0.009)	(0.009)	(0.018)			
Trade in 1992				0.193^a	0.193^a	0.057^a
				(0.009)	(0.009)	(0.020)
$\theta_s \times$ Pop-weighted Distance Groups	Y	Y	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y	Y	Y
Observations	1779937	1779937	1779937	1779937	1779937	1779937
# of Destination-Origin groups	2926	2926	2926	2926	2926	2926
1st-Stage F-stat		36.22			36.22	
R-squared	0.711		0.721	0.708		0.721

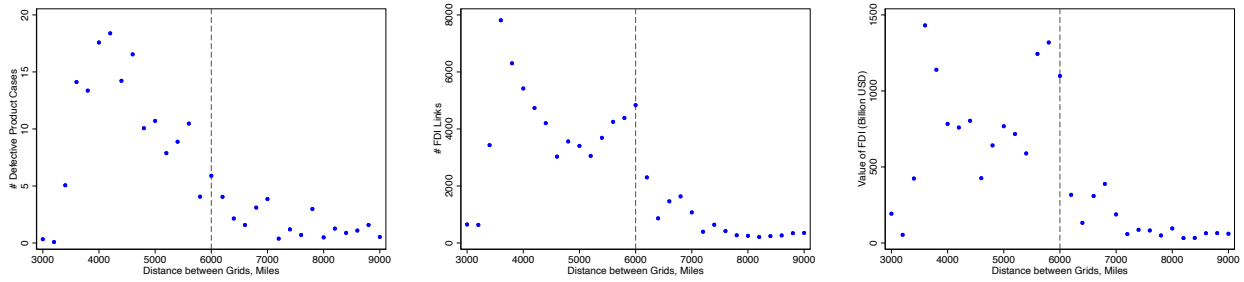
Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10. θ_s is from Caliendo and Parro (2015). Pop-weighted distance groups refer to the six population-weighted distance dummy variables. The trade in 1989 and 1992 was also at the destination-origin-HS6 sector level. To prevent losing too many observations due to zero historical trade flows, we apply an inverse hyperbolic sine transformation to the trade flows in 1989 and 1992.

In Online Appendix B.4, we provide a host of additional robustness checks. These include: (i) excluding products typically transported by air (the share of air shipment is higher than 5% or 1% in total US import value of the HS6 product) to isolate the communication channel from potential transportation cost effects; (ii) using alternative estimates of trade elasticity θ_s from Giri et al. (2020) and specifications without sector-specific elasticities (implicitly assuming common θ_s across sectors); (iii) employing narrower distance windows when constructing the key instrumental variable z_{od} and adopting multiple instrumental variables to take into account the spillover effects to shorter-distance flights between two countries; (iv) varying the radius used to define airport catchment area from 100 to 500 miles, both in great circle distances and in "normalized road distance" as in Allen and Arkolakis (2014); (v) altering the minimum frequency threshold when defining direct flight connections from weekly to twice-weekly and daily flights; (vi) eliminating borders between countries within the Schengen Zone to make it easier for those countries to connect to nearby airports. Across all specifications, we consistently find positive and statistically significant effects of air connectedness on trade flows, confirming the robustness of our main results.

5.2 Impact of Air Connectedness on the Likelihood of Product Defects

In this section, we analyze the impact of distance and direct flights on the number of defective-product lawsuits and foreign ownership links at the grid-pair level. Before we perform regression discontinuity (RD) analysis, we first examine the relationship between distance and the outcome variables graphically. In panel (a) of Figure 2, we observe that the number of defective-product cases between two grids declines with their distance, and the decline is smooth at around 6,000 miles. In contrast, panels (b) and (c) show that the number of foreign ownership links and the total revenue of foreign subsidiaries have a clear discontinuity at 6,000 miles. Combined with the fact that the 6,000-mile threshold significantly reduces direct flights, this suggests that direct flights have a significant impact on foreign ownership links but not on business lawsuits. Since business lawsuits are proxies for product defects, and foreign ownership links are a proxy for business activities, the patterns in Figure 2 suggest that direct flights reduce the probability of product defects.

Figure 2: Number of Litigation and foreign ownership links between Grid Pairs, by Distance



(a) # defective product cases (b) # foreign ownership links (c) Subsidiary revenue

Notes. We group grid pairs into 200-mile bins. Each dot represents the number of cases, the number of foreign ownership links, or the total revenue of foreign subsidiaries (in billion USD) in the preceding 200 miles. All statistics are calculated using a common set of grid pairs, where the grids correspond to the 1,639 cities in the selected top 100 trading countries with airports in the TFS data. After dropping all the grid pairs within the same country, we have 2,582,388 potentially connected grid pairs. Grid pairs with no lawsuits or foreign ownership links are kept in the sample and assigned a number of zero.

We also conduct standard RD diagnostics. Online Appendix B.5 shows that the density of grid pairs and predetermined economic and geographic covariates are continuous at the 6,000-mile cutoff. In addition, Online Appendix Figure B-1 confirms that the discontinuity in connectedness at 6,000 miles is present in the study period but not in the pre-period.

5.2.1 Grid-pair Level Regression Discontinuity

In Table 6, we report the results of the sharp RD analysis. The first four columns focus on the impact of distance on the number of defective-product lawsuits. Columns (1)-(3) present

the estimates from equation (11), allowing for linear trends on either side of the 6,000-mile threshold and experimenting with different bandwidths: 500 miles, 800 miles, and the optimal bandwidth, calculated according to Calonico, Cattaneo and Titiunik (2014). Column (4) displays the results of the RD regression using second-order polynomial functions. Having a distance of less than 6,000 miles has almost no effect on the number of lawsuits: the point estimates are all insignificantly different from zero.

In contrast, columns (5)-(8) show that the number of foreign ownership links increases when the distance between two grids is less than 6,000 miles, consistent with the findings in Campante and Yanagizawa-Drott (2018); the estimates are statistically significant at the 800-mile and MSE-optimal bandwidths. Given the average number of foreign ownership links is around 0.04 (see the last row of the table), having a distance below 6,000 miles increases the number of ownership links between two grids by 65% to 87%, depending on the specification. Even if we take the insignificant estimates for the defective-product cases at face value, the same treatment increases the number of cases by around 17% of the mean in column (1), compared with 65% to 87% of the mean for foreign ownership links. Therefore, the impact of distance on bilateral business activities is much larger than the impact on business lawsuits, which is consistent with our hypothesis that direct flights reduce the *probability* of product defects.

Table 6: Sharp RD: Effects of Distance on Number of Cases and Foreign Ownership Links at the Grid-pair Level

Dep. Var.:	# DP Cases$_{ji}$ (×10)				# FDI Links$_{ji}$			
Bandwidth (Miles):	500	800	Optimal		500	800	Optimal	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}(d_{ji} \leq 6K$ miles)	0.0001	0.0000	0.0000	-0.0000	0.0272	0.0321^b	0.0303^c	0.0327^c
	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0215)	(0.0154)	(0.0168)	(0.0186)
Observations	2582388	2582388	2582388	2582388	2582388	2582388	2582388	2582388
Observations, effective	355730	561598	558014	629118	355730	561598	499696	867664
Polynomial Order	1	1	1	2	1	1	1	2
Kernel Choice	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Bandwidth	500	800	794	899	500	800	708	1259
Mean of Dep. Var.	0.0006	0.0007	0.0007	0.0008	0.0421	0.0417	0.0420	0.0377

Notes. Local polynomial regression discontinuity estimates. The dependent variables are the number of "defective product" business lawsuits and foreign ownership links. When the dependent variable is the number of litigation cases, the values are scaled up 10 times due to the small number of business lawsuit cases. The running variable is the distance between grids, discontinuity at 6,000 miles. The optimal bandwidth is selected using the MSE-optimal bandwidth selector, and the triangular kernel function is adopted (both are default choices in Stata's `rdrobust` command). "Observations" refer to the total number of grid pairs across all distance ranges, while "effective observations" only include grid pairs within the bandwidth. Robust standard errors are in parentheses. Significance levels: a: 0.01, b: 0.05, c: 0.10.

Next, we present the fuzzy RD results in Table 7. We instrument the average connect-

edness between two grids, $\bar{\kappa}_{ji}$, with the indicator of distance below 6,000 miles, controlling for linear or quadratic functions on both sides of the threshold. The lower panel of the table displays the first-stage results, confirming that having a distance of less than 6,000 miles significantly increases the probability of being connected by direct flights. Except for column (4), the first-stage Kleibergen-Paap F-statistics are all greater than 10. The second-stage results show that direct flights do not increase the number of lawsuits (columns (1)-(4)), but they increase the number of foreign ownership links, significantly so at the 800-mile and linear optimal bandwidths (columns (6)-(7)). The overall results are consistent with the sharp RD results and support our hypothesis that direct flights reduce the probability of business disputes between firms.

Table 7: Fuzzy RD: Effects of Direct Flights on Number of Cases and Foreign Ownership Links at the Grid-pair Level

Dep. Var.:	# DP Cases$_{ji}$ (×10)				# FDI Links$_{ji}$			
Bandwidth (Miles):	500	800	Optimal		500	800	Optimal	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$AvgConnected_{ji}$	0.011	0.000	0.009	0.048	2.887	1.673^b	1.669^b	2.190
	(0.025)	(0.015)	(0.025)	(0.081)	(2.291)	(0.800)	(0.789)	(1.911)
First Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}(d_{ji} \leq 6K$ miles)	0.009^a	0.019^a	0.010^a	0.003^b	0.009^a	0.019^a	0.019^a	0.011^a
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations, total	2582388	2582388	2582388	2582388	2582388	2582388	2582388	2582388
Observations, effective	355730	561598	362540	522270	355730	561598	566102	691434
Polynomial Order	1	1	1	2	1	1	1	2
Kernel Choice	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Bandwidth	500	800	510	742	500	800	807	990
1st-Stage F-stat	74.6	492.9	81.2	3.9	74.6	492.9	504.1	101.2
Mean of Dep. Var.	0.0006	0.0007	0.0006	0.0007	0.0421	0.0417	0.0416	0.0397

Notes. Local polynomial regression discontinuity estimates. The dependent variables are the number of business lawsuits containing the keyword "defective product" and the number of foreign ownership links. When the dependent variable is the number of litigation cases, it is scaled up 10 times for better presentation. The endogenous regressor is $\bar{\kappa}_{ji}$, the average connectedness between the two grids over 2011–2017, consistent with the time span of the business lawsuits and foreign ownership links. The lower panel reports the corresponding first-stage regressions. The running variable is the distance between grids, discontinuity at 6,000 miles. The optimal bandwidth is selected using the MSE-optimal bandwidth selector, and the triangular kernel function is adopted (both are default choices in Stata's `rdrobust` command). "Observations" refer to the total number of grid pairs across all distance ranges, while "effective observations" only include grid pairs within the bandwidth. Robust standard errors are in parentheses. Significance levels: a: 0.01, b: 0.05, c: 0.10.

We perform several robustness checks in Online Appendix B.6. Our results are robust to: (i) using a uniform instead of a triangular kernel (Table B-16); (ii) excluding grid pairs involving the United States, the country that engages in the largest number of business lawsuits and foreign ownership links in our sample (Table B-17); (iii) controlling for the

absolute differences in population, GDP per capita, and coordinates between the two grids (Table B-18); (iv) varying the bandwidth from 200 to 3,000 miles (Figure B-6); and (v) transforming the dependent variables with the inverse hyperbolic sine or $\log(1+y)$ to account for the many zeros (Tables B-19 and B-20). Across these specifications, the effect of distance on foreign ownership links is consistently larger than that on defective-product lawsuits, and the latter is never significantly different from zero.

5.2.2 Country-pair Level Evidence

Both the number of litigation cases and the number of foreign ownership links contain a large number of zeros at the grid-pair level, which makes it infeasible to control for, or divide by, the number of foreign ownership links in the grid-pair regressions. Aggregating the data to the country-pair level reduces the density of zeros, so we complement the RD analysis with country-pair level regressions.

Table 8 examines how the ratio of defective-product cases to foreign ownership links—a direct measure of the probability of product defects within a given amount of foreign investment—responds to the distance-discontinuity index z_{od} used in our country-pair trade regressions. The first four columns report OLS estimates in logs; because the log specification drops country pairs in which the ratio is zero, the last four columns re-estimate the relationship with Poisson Pseudo Maximum Likelihood (PPML), which retains the zeros. The estimates are negative in all eight columns. The OLS estimates are significant without the population-weighted distance-group controls and lose precision once these controls are added; the PPML estimates are significant in the specifications with country fixed effects, including the most demanding one that also includes the distance-group controls. The country-pair evidence thus supports the grid-pair finding: the reduction in communication costs associated with distances below 6,000 miles lowers the probability of problematic transactions.

Table 8: Effects of Distance Discontinuity on Ratio of Cases to Foreign Ownership Links

Dep. Var.:	log(# DP Cases$_{od}$ / FDI Links$_{od}$)				# DP Cases$_{od}$ / FDI Links$_{od}$			
	OLS				PPML			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\sum_{ij\in od} w_{ij}\mathbb{1}\left(d_{ij}\leq 6K \mid d_{ij}\in(5.5K, 6.5K]\right)$ miles)	-1.250^{a}	-1.359^{a}	-0.811	-0.779	-0.287	-0.670^{c}	-0.911	-1.127^{b}
	(0.410)	(0.396)	(0.672)	(0.550)	(0.920)	(0.374)	(1.047)	(0.470)
Control: Pop-weighted Distance Groups	N	N	Y	Y	N	N	Y	Y
Origin-Country FE	N	Y	N	Y	N	Y	N	Y
Dest.-Country FE	N	Y	N	Y	N	Y	N	Y
Observations	193	157	193	157	855	566	855	566
# of Destination-Origin groups	193	157	193	157	855	566	855	566
R-squared	0.038	0.725	0.118	0.753				
pseudo-R-squared					0.001	0.405	0.044	0.407

Notes. The dependent variables are the ratio of defective product lawsuits to foreign ownership links at the country-pair level. The first four columns are logarithm OLS regressions and the last four columns are PPML regressions. The key regressor is the pop-weighted share of grid pairs within a country pair whose distance falls just below 6,000 miles (within the 5,500-6,500 mile window), which is the same as our country-pair level main regressions. Standard errors are clustered at the country-pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10.

We also directly control for the number of foreign ownership links, as an alternative to using it as the denominator. Online Appendix Tables B-21 and B-22 regress the number of defective-product cases on z_{od}, controlling for FDI links either as a continuous variable or through category indicators. The OLS coefficients are negative throughout and significant without the distance-group controls, while the PPML estimates are noisier and insignificantly different from zero. We prefer the ratio specification in Table 8, which maps directly into a failure probability, and present the controlling specifications as robustness checks.

6 Counterfactual Analysis

In this section, we conduct counterfactual analysis with respect to air connectedness κ_{od} and thus communication costs m_{od}. We focus on three episodes: the improvement in air connectedness from 2002 to 2011, the sudden decline from 2019 to 2020 due to the COVID-19 pandemic, and the closure of Russian and Ukrainian airspace after February 2022.

6.1 General Equilibrium

In this section, we use the "exact hat algebra" method introduced by Dekle et al. (2007) to calculate the welfare gain from the establishment of international direct flights. We first discuss the equilibrium conditions and close the model introduced in Section 2. The price index of sector s composite goods in the destination country d (see equation (6)) depends

on the unit cost of the input bundle, c_{os}. The latter can be written as

$$c_{os} = \Upsilon_{os} w_o^{a_{L,os}} \prod_k P_{ok}^{a_{ok,s}}. \tag{12}$$

$a_{L,os}$ is the value added share in sector s, country o, and $a_{ok,s}$ is the share of composite input from sector k in total production costs of sector s in country o. Υ_{os} is a constant that equals $a_{L,os}^{-a_{L,os}} \prod_k a_{ok,s}^{-a_{ok,s}}$.

Consumers in each country have Cobb-Douglas preferences over composite goods of each sector. We denote the consumption shares by ξ_{os}. Denoting the total sales of sector s in country o as R_{os}, the goods market clearing condition can be written as

$$R_{os} = \sum_d \pi_{ods} \left(\xi_{ds} E_d + \sum_k a_{ds,k} R_{dk} \right), \tag{13}$$

where E_d is the total expenditure in the destination country d, and π_{ods} is defined as the trade shares

$$\pi_{ods} \equiv \frac{X_{ods}}{X_{ds}} = \frac{T_{os} \left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od} \right)^{-\theta_s}}{\sum_o T_{os} \left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od} \right)^{-\theta_s}}. \tag{14}$$

The total expenditure, in turn, equals total income plus an exogenous transfer that captures trade deficits:

$$E_d = w_d L_d + D_d.$$

We impose $\sum_d D_d = 0$ so that consumers will consume all final goods produced, and Walras's Law holds.

Equation (13) implies that all composite goods are consumed by domestic or foreign consumers or used as intermediate inputs by downstream sectors. We can solve R_{os} from the $N \times S$ linear equations like (13) given shares, prices, and exogenous transfers.

Finally, in equilibrium, labor markets clear in all countries. We express this condition with wage bills

$$w_o L_o = \sum_s a_{L,os} R_{os}. \tag{15}$$

In our counterfactual analysis, we consider a change in the exogenous communication cost from m_{od} to m'_{od}. Since we have specified earlier that $m_{od} = e^{\beta_0 - \beta \kappa_{od}}$, the constant β_0 cancels and the change in communication costs equals $\hat{m}_{od} \equiv m'_{od}/m_{od} = e^{-\beta\left(\kappa'_{od} - \kappa_{od}\right)}$, where κ_{od} and κ'_{od} are the connectedness index between countries o and d in the base and

later periods, respectively.[17] Using the exact hat algebra method, we can solve the change in equilibrium variables as functions of the changes $\hat{m}_{od}$, the base period shares, and known parameters θ_s and ρ_s. We relegate the detailed formulas and algorithms to Online Appendix C.

One caveat applies to the interpretation of all the counterfactual exercises in this section. New direct flights are not free, but the resources they require depend on their source: routes made possible by aviation agreements or aircraft technology involve little additional social cost, and airlines operate routes at a profit, whereas government investment in aviation infrastructure is a real resource cost. Our counterfactuals take the observed changes in connectedness as given and quantify the communication benefits; a full welfare accounting would add airline profits and subtract public infrastructure costs, neither of which we quantify.

6.2 Improvement in Air Connectedness from 2002 to 2011

In our quantitative applications, we focus on the set of countries in the World Input-Output Database (WIOD, Release 2013) and use 2002 as the base year. The first counterfactual ends in 2011, the last year covered by the WIOD 2013 release, so that the model's predictions can be benchmarked against the observed changes in trade shares over the same window. We focus on 39 countries with non-missing data on bilateral air connectedness κ_{od}. The set of countries is listed in Table 9. While the estimation in Section 5 uses only the 2016 cross-section, the TFS data allow us to construct the connection index year by year; the counterfactuals in this section use this panel to measure observed changes in connectedness, which we translate into changes in communication costs using the estimated β.

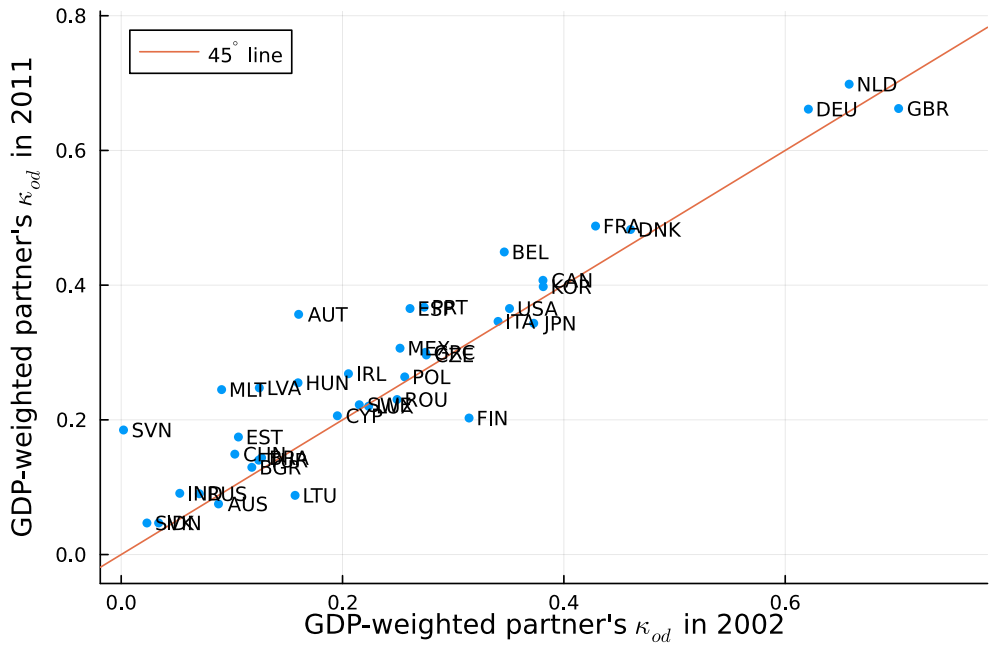
Figure 3 compares air connectedness in 2011 with that in 2002. For each country o, we calculate the average of κ_{od} across all its partners, d, weighting each partner by its base-period GDP, in the base year 2002 and the later year 2011.[18] In most countries, connectedness with foreign countries improved during this period. European countries, such as Austria, Slovenia, Malta and Latvia, saw large increases in κ_{od} (between 0.11 and 0.17). Meanwhile, the air

[17] We hold fixed the fundamental productivities T_{os}, the labor endowments L_o, the iceberg trade costs t_{od}, the base-year input-output structure and expenditure shares, and the sectoral elasticities θ_s and ρ_s. Wages, prices, expenditures, gross output, and the bilateral trade shares in both final and intermediate use all adjust endogenously. We normalized the US wage to one. The exogenous transfers scale with income in the counterfactual equilibrium, $D'_d = \hat{w}_d D_d$, one of the two conventions on trade imbalances in static trade models discussed by Costinot and Rodríguez-Clare (2014); total expenditure then changes in proportion to the wage, $\hat{E}_d = \hat{w}_d$, so the change in the real wage coincides with the change in real absorption. Online Appendix D.3 shows that the results are nearly identical when D_d is instead held fixed in numeraire units.

[18] Throughout this section, country-level connectedness refers to this GDP-weighted average of κ_{od} across partner countries. The weights are the countries' total value added in 2002 from the WIOD.

connectedness of the United States increased by only 0.01. Using our empirical estimate $\beta = 0.026$, we use equation (C-1) to predict the change in communication costs. We then solve the change in equilibrium variables following the exact hat algebra method outlined in the previous section, combining base period shares and known parameters θ_s and ρ_s.[19]

Figure 3: Average κ_{od} in 2011 v.s. 2002



Notes. For each country, the figure plots the average of κ_{od} across its partner countries in 2011 against the same average in 2002. The average weights each partner by its base-period GDP, measured as total value added in 2002 from the WIOD. The sample is the 39 countries listed in Table 9.

In columns (4)–(7) of Table 9, we present percentage changes in equilibrium outcomes due to the change in air connectedness from 2002 to 2011. We report the changes in domestic expenditure shares in final consumption, intermediate input usage and total absorption, as well as the changes in real wages. The changes in domestic expenditure shares in final consumption and intermediate inputs are not the same, but they are highly correlated; drops in these shares mean that countries spend more on foreign goods and that consumer welfare improves. Online Appendix Figure D-1 examines the relationships between the changes in connectedness, domestic expenditure shares and real wages in detail. Consistent with the changes in air connectedness, Malta has the highest gains from the improvement in air connectedness during this period (1.42%). Despite not having the highest improvement in air connectedness, Slovenia ranks second in the increase of real wages (0.89%). This is because

[19]We aggregate product-level ρ_s to the 35 sectors in WIOD and map the estimates of θ_s in Caliendo and Parro (2015) to the same sectors in WIOD. For non-tradable sectors, we assume their ρ_s and θ_s are the median of tradable sectors.

they started from very low levels of κ_{od} in 2002 and are small economies that rely on trade. In Online Appendix D.1, we show that the welfare changes in this counterfactual exercise are robust to replacing the baseline trade elasticities θ_s from Caliendo and Parro (2015) with those from Giri et al. (2020), or with a common θ across sectors. In Online Appendix D.2, as a conservative robustness check, we further show that one third to two thirds of the GDP-weighted welfare gain remains even when the communication-cost shock is confined to goods sectors that are not intensively shipped by air.

How important is direct flight connectedness in explaining the observed changes in trade shares and welfare over this period? For trade shares, the question can be answered directly: we compare the changes in trade shares implied by the model with the changes observed in the WIOD over the same window, using the foreign share of total absorption as the summary measure. Online Appendix Table D-1 reports the comparison country by country. For the world as a whole, the foreign share of absorption rises by 2.0 percentage points in the data, from a 2002 level of 10.3%, while the model predicts an increase of 0.10 percentage points: the improvement in air connectedness explains 4.9% of the observed increase. The share is considerably higher for the countries with the largest improvements in connectedness: 48% in Austria and 44% in Belgium.

Answering the same question for welfare requires one further step, because the welfare gains associated with the observed changes are not directly observed. We assume that changes in sectoral trade costs are the only driver of the observed changes in trade shares. Under this assumption, the sufficient-statistics formula of Costinot and Rodríguez-Clare (2014)—which combines the observed trade shares in the two periods, the trade elasticities and the input-output coefficients—delivers the welfare gains implied by the observed shifts. We then compare the welfare changes predicted by our counterfactual with this benchmark. For the world as a whole, the observed shifts imply a welfare gain of 1.49%, of which the model predicts 0.03%: the improvement in air connectedness accounts for 2.1% of the implied welfare gains. The share is again considerably higher for the countries with the largest improvements in connectedness: 41% for Austria and 14% for Belgium.

Table 9: Impact of the change in air connectedness from 2002 to 2011

Country	κ_o^{2002}	$\Delta\kappa_o$	$\Delta\% \, s_{dom}^{final}$	$\Delta\% \, s_{dom}^{input}$	$\Delta\% \, s_{dom}$	$\Delta\% \, w/P$
Australia	0.09	-0.01	0.00	-0.02	-0.01	-0.01
Austria	0.16	0.17	-0.75	-2.00	-1.30	0.38
Belgium	0.35	0.09	-0.50	-2.46	-1.56	0.41
Bulgaria	0.12	0.01	-0.07	-0.18	-0.12	0.06
Brazil	0.13	0.01	-0.01	-0.04	-0.02	0.01
Canada	0.38	0.02	-0.03	-0.07	-0.05	0.02
China	0.10	0.04	-0.04	-0.09	-0.07	0.03
Cyprus	0.20	0.01	-0.01	-0.20	-0.08	0.04
Czechia	0.28	0.02	-0.08	-0.20	-0.15	0.06
Germany	0.62	0.03	-0.08	-0.14	-0.11	0.04
Denmark	0.46	0.02	0.02	0.11	0.06	-0.01
Spain	0.26	0.09	-0.36	-0.78	-0.56	0.16
Estonia	0.11	0.06	-0.46	-0.97	-0.73	0.23
Finland	0.31	-0.10	0.11	0.32	0.22	-0.07
France	0.43	0.05	-0.28	-0.49	-0.37	0.10
United Kingdom	0.70	-0.03	0.01	0.05	0.03	-0.01
Greece	0.27	0.02	-0.04	-0.11	-0.06	0.02
Hungary	0.16	0.08	-0.68	-1.87	-1.25	0.28
Indonesia	0.03	0.01	-0.05	-0.19	-0.12	0.03
India	0.05	0.03	-0.02	-0.05	-0.03	0.01
Ireland	0.21	0.06	-0.09	-0.61	-0.35	0.12
Italy	0.34	0.00	-0.38	-0.54	-0.46	0.13
Japan	0.37	-0.02	0.01	0.01	0.01	-0.01
South Korea	0.38	0.01	-0.03	-0.16	-0.10	0.02
Lithuania	0.16	-0.06	-0.51	0.87	0.06	-0.21
Luxembourg	0.22	-0.00	0.02	0.64	0.39	-0.03
Latvia	0.12	0.11	-0.49	-0.64	-0.56	0.30
Mexico	0.25	0.05	-0.06	-0.24	-0.12	0.03
Malta	0.09	0.14	-1.78	-5.34	-3.24	1.42
Netherlands	0.66	0.04	-0.14	-0.34	-0.22	0.09
Poland	0.26	0.01	-0.12	-0.36	-0.23	0.07
Portugal	0.27	0.08	-0.45	-1.09	-0.73	0.24
Romania	0.25	-0.02	-0.28	-0.91	-0.57	0.22
Russia	0.07	0.02	-0.01	-0.01	-0.01	0.00
Slovakia	0.02	0.02	-0.66	-1.21	-0.96	0.18
Slovenia	0.00	0.16	-0.87	-3.09	-1.96	0.89
Sweden	0.22	0.01	0.00	0.05	0.03	-0.00
Turkey	0.12	0.01	-0.09	-0.20	-0.14	0.06
USA	0.35	0.01	-0.01	-0.01	-0.01	0.00
Weighted Average	0.36	0.01				0.03

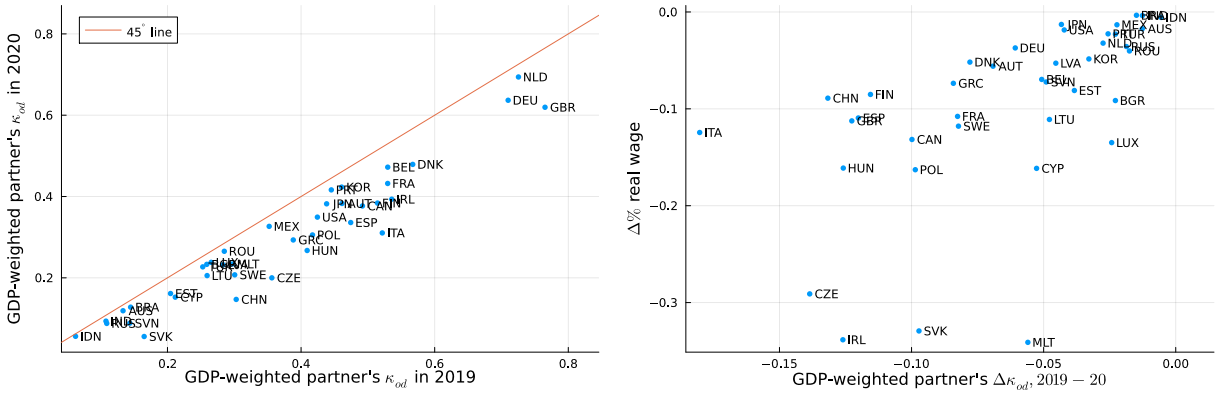
Notes. This table reports the impact of the change in air connectedness from 2002 to 2011. κ_o^{2002} and $\Delta\kappa_o$ are the level in 2002 and the change from 2002 to 2011 of country o's connectedness with its partners, averaged across partners using their base-period GDP as the weights. $s_{dom}^{final}, s_{dom}^{input}$ and s_{dom} denote the domestic expenditure shares in final consumption, intermediate input usage and total absorption. w/P indicates real wage, in which the price index is calculated as $P_d = \prod_s P_{ds}^{\xi_{ds}}$. The last row reports the weighted average of all countries, using the base-period GDP as the weights.

6.3 The Drop in Air Connectedness from 2019 to 2020

The COVID-19 pandemic caused disruptions to international travel. As shown in Panel (a) of Figure 4, air connectedness with partner countries decreased from 2019 to 2020 for all countries in our sample. Italy had the largest drop in air connectedness, with the GDP-weighted average of κ_{od} across partner countries falling by 0.18, followed by the Czech Republic (0.14), China (0.13), Ireland and Hungary (0.13 each). We evaluate the impact of such declines using the same approach as in the previous section.[20]

Online Appendix Table D-2 reports the predicted changes in domestic expenditure shares and real wages in each country. Ireland and Malta, two small open economies with large drops in connectedness, experienced the largest declines in real wages (0.34% each). Our counterfactual analysis also predicts large declines in real wages in two Eastern European countries, Slovakia and the Czech Republic, which experienced large declines in air connectedness and are predicted to have large increases in the share of domestic expenditures. Panel (b) of Figure 4 summarizes the country-level welfare effects, plotting the change in real wages against the GDP-weighted change in partners' connectedness.

Figure 4: The drop in air connectedness from 2019 to 2020 and its welfare effects



(a) Average κ_{od} in 2020 v.s. 2019 (b) Change in real wages and partner's $\Delta\kappa_{od}$

Notes. In both panels, averages across partner countries weight each partner by its base-period GDP, measured as total value added in 2002 from the WIOD. Panel (a) plots the average of κ_{od} across partner countries in 2020 against the same average in 2019. Panel (b) plots the model-implied percentage change in real wages against the change in the average of κ_{od} across partner countries from 2019 to 2020.

[20]Due to data limitations, we do not have sectoral trade shares and input-output tables for 2019. We therefore use the 2002 trade shares in the hat algebra and feed in the change in air connectedness from 2019 to 2020, so the exercise measures the effect of the decline in connectedness on an economy with the trade structure of 2002.

6.4 The Closure of Russian and Ukrainian Airspace

In February 2022, Ukrainian airspace was closed to civil aviation, more than thirty countries closed their airspace to Russian carriers, and Russia barred the carriers of those countries in return. The episode changes air connectedness through two channels. First, direct flights between Russia and the banning countries stop altogether; we refer to this channel as the *direct effect*. Second, flights between countries whose original flight paths cross the closed airspace must fly around it, which lengthens the route; we refer to this channel as the *detour effect*. Both changes tend to lower connectedness between countries.

To predict the flight detour after the airspace closure, we treat the union of the Russian and Ukrainian polygons, with a 0.5° buffer, as a no-fly zone. We overlay a $1° \times 1°$ grid on the globe, so that the world map is divided into small cells and the no-fly zone is shown as an obstacle. Each cell is either open or blocked, where a cell is blocked if it falls inside the no-fly zone. We identify all grid cells containing airports, and for each pair of airport-grids whose great-circle route passes through the obstacle, we search for the shortest path that connects the two grids while stepping only through open cells.[21]

Table 10 compares the resulting detour distances with the great-circle distances. The rows indicate the group of the great circle distance and the columns indicate the group of the detour distance, with ∞ indicating that direct flights are banned because one of the grid falls inside the closed airspace. Among the 660,268 airport-grid pairs whose great-circle path crosses the closed airspace, 145,990 (22.1%) have an endpoint inside the closed airspace and lose the possibility of a direct connection entirely.[22] For the remaining pairs, the distance distribution shifts up substantially. For example, of the 57,512 affected pairs whose great-circle distance lies in $(5{,}500, 6{,}000]$ miles, only 20,270 (35%) remain in that band, while 45% are pushed above 6,000 miles and a further 20% lose feasibility.

[21] We use the A* algorithm (Hart, Nilsson and Raphael, 1968), which is similar to Dijkstra's algorithm used in Besedeš et al. (2024) but more efficient: because the destination is known, A* can direct its search toward the target rather than exploring in all directions, reaching the same optimal solution while visiting far fewer cells. Because each step moves between adjacent grid cells, the raw path has a staircase shape; we smooth it with greedy shortcutting, which replaces the zig-zag segments with direct great-circle arcs without increasing the total distance. Online Appendix E gives the algorithm in full.

[22] We exempt pairs involving countries that continued to operate direct flights with Russia after February 2022, which we identify from Aeroflot's post-2022 destination list. These pairs neither reroute nor lose their connection, which is why Russia retains some connectedness in the counterfactual.

Table 10: Great-circle vs. detour distance groups after the closure of Russian and Ukrainian airspace

GC \ Detour	0–1K	1–2K	2–4K	4–5.5K	5.5–6K	6–6.5K	6.5–8K	>8K	∞	Row Total
0–1K	5,806	265	1	0	0	0	0	0	6,776	12,848
1–2K	0	19,257	2,340	93	0	0	0	0	23,179	44,869
2–4K	0	0	61,450	6,861	542	611	198	0	32,750	102,412
4–5.5K	0	0	0	58,696	16,766	11,542	8,646	24	34,138	129,812
5.5–6K	0	0	0	0	20,270	13,985	11,521	216	11,520	57,512
6–6.5K	0	0	0	0	0	21,576	22,235	880	9,254	53,945
6.5–8K	0	0	0	0	0	0	88,008	28,260	17,607	133,875
>8K	0	0	0	0	0	0	0	114,229	10,766	124,995
Total										660,268

Notes. Rows group grid pairs by their great-circle distance; columns group the same pairs by their shortest detour distance avoiding Russian and Ukrainian airspace. Cells report the number of grid pairs. The ∞ column collects pairs with an endpoint inside the closed airspace, for which no feasible path exists. Pairs are counted once, so that A–B and B–A are a single observation. Distances are in miles and are computed on a sphere.

We map these distances into a counterfactual connectedness index in three steps. First, we estimate a linear probability model of whether a grid pair has a weekly direct flight on indicators for great-circle distance groups using the 2019 cross-section and predict the change in the probability of airports being connected. Column (1) of Online Appendix Table E-1 reports the estimates we use.[23] Given the estimated equation and the change in the distance groups in Table 10, we predict the change in the probability between airport grids i' and j', $\Delta\text{Prob}_{j'i'}$, with the probability after the detour being zero for pairs with one end inside the no-fly zone. Second, for all grid pairs, including those without airports, we calculate the change in connectedness as the average of $\Delta\text{Prob}_{j'i'}$ in the airport grids within the 200-mile neighbourhoods, i.e., $\Delta\kappa_{ji} = \overline{\Delta\text{Prob}_{j'i'}}$. Finally, since the country-pair index aggregates grid pairs linearly (equation (8)), the grid-level changes aggregate directly to country-pair changes $\Delta\kappa_{od}$, without constructing any counterfactual level of the index, and we feed the implied change in communication costs $\hat{m}_{od} = e^{-\beta\Delta\kappa_{od}}$, again with $\beta = 0.026$, into the exact hat algebra.[24]

Across the 1,482 WIOD country pairs, the connection index falls on average by 0.008, about 1.8% of its 2019 level. The overall change is small, and it is driven mainly by Russia losing its connections with other countries. The first two columns of Table 11 report air

[23] This mapping does not rely on the 6,000-mile discontinuity that underlies our IV. We use the discontinuity as an identification device for the impact of direct flight on m_{od}. However, it only focuses on the change from below 6000 miles to above. As Table 10 shows, many detours move grid pairs across distance groups entirely above or below 6,000 miles, and these pairs would be assigned no change in connectedness if we only focus on the discontinuity at 6,000 miles. We therefore use the full profile of connection probabilities across distance groups.

[24] As in Section 6.3, the hat algebra uses the 2002 trade shares and input-output tables, since we do not have trade shares for 2019 (footnote 20).

connectedness before the closure and its change. Russia is affected far more than any other country: its average connectedness with partner countries falls from 0.10 to 0.01, and the small residual reflects the partners that did not ban Russian carriers, notably China, India and Turkey, which retain their direct flights. Every other country in the sample loses at most 0.005, and the largest declines are in Germany, Austria, Lithuania, the Czech Republic and Finland.

How much of the decline comes from the direct effect, and how much from the detour effect? The third column of Table 11 reports, for each country, the share of its decline in connectedness that is attributable to pairs with Russia. The split is heterogeneous: the direct effect accounts for over 90% of the decline in Slovenia and South Korea, but less than half in the countries whose own routes to East Asia crossed the closed airspace, such as Austria, Denmark, Poland.[25] The countries where the detour effect looms large, however, have very small overall declines in connectedness (at most 0.005). The reason lies in how a detour enters our index: a rerouted route is assigned the change in the probability of a weekly direct flight between its old and its new distance group, and because only a small fraction of airport pairs at any distance operate a direct flight, this change is a fraction of a percentage point. The welfare results below are therefore dominated by the direct effect. The detour effect does lengthen flights substantially on the routes that remain,[26] which points to a caveat of our quantification: our estimate of β captures the effect of having a direct flight, not of its duration, so the longer flights raise the time cost of face-to-face communication, a margin that our welfare numbers miss. The welfare losses reported below should therefore be read as conservative, particularly for the countries whose routes to East Asia became substantially longer.

The last two columns of Table 11 report the welfare consequences. Russia loses the most, with real wages falling 0.14%—about three times the GDP-weighted average welfare loss in the COVID-19 counterfactual—and its domestic expenditure share rising 0.36%. Outside Russia the losses are an order of magnitude smaller but not uniform: Lithuania loses 0.08% and Latvia 0.03%, more than countries with comparable declines in connectedness. Both are small open economies; Lithuania, in particular, sources about one third of its imported intermediate inputs from Russia in the base year, the largest share in the sample.

[25] The share is exactly zero for Slovakia, Canada, Luxembourg, Mexico, Brazil and Australia: although these countries participate in the mutual ban, they had no direct-flight connectedness with Russia in the 2019 data, so their entire decline in connectedness comes from detours.

[26] Columns (4) and (5) of Table 11 report each country's GDP-weighted great-circle distance to its partners and the average increase in flight distance on its rerouted routes. For the countries on the Europe–Northeast Asia corridors the increases are substantial: rerouted routes lengthen on average by about 660 miles for Germany, 830 for the United Kingdom, 1,000 for Denmark and almost 1,200 for Sweden.

Table 11: Impact of the closure of Russian and Ukrainian airspace

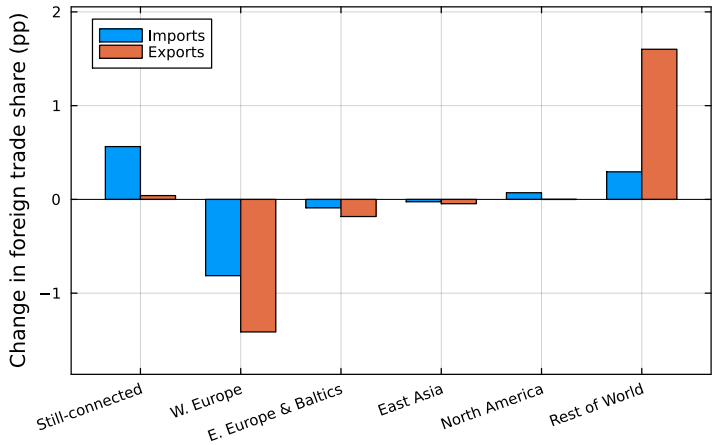
Country	κ_o^{2019}	$\Delta\kappa_o$	% RUS	d_{gc}	$\Delta d^{\text{affected}}$	$\Delta\% \, s_{dom}$	$\Delta\% \, w/P$
Russia	0.103	-0.092	0.0	4173	∞	0.357	-0.140
Germany	0.665	-0.005	54.3	3914	662.4	0.029	-0.009
Austria	0.429	-0.005	47.9	3824	442.8	0.032	-0.011
Lithuania	0.236	-0.005	31.3	3784	309.1	0.244	-0.085
Czechia	0.327	-0.005	70.6	3775	548.2	0.061	-0.024
Finland	0.467	-0.005	75.4	3745	624.6	0.086	-0.026
Poland	0.384	-0.004	23.4	3800	292.6	0.024	-0.007
Latvia	0.260	-0.004	54.9	3771	745.1	0.064	-0.031
United Kingdom	0.733	-0.004	53.1	3800	830.2	0.009	-0.002
Bulgaria	0.235	-0.004	55.5	4083	70.8	0.093	-0.023
Denmark	0.520	-0.003	46.4	3661	1003.0	0.019	-0.002
Italy	0.481	-0.003	61.5	4062	328.8	0.013	-0.005
Belgium	0.495	-0.003	66.5	3685	712.5	0.010	-0.004
Malta	0.271	-0.003	67.8	4214	142.9	0.023	-0.008
Hungary	0.378	-0.003	67.3	3881	154.0	0.025	-0.022
Netherlands	0.674	-0.003	66.4	3689	741.9	0.021	-0.004
Spain	0.438	-0.003	58.8	4045	444.5	0.009	-0.003
Greece	0.356	-0.003	66.2	4210	50.0	0.018	-0.009
Cyprus	0.197	-0.003	66.6	4466	NA	0.081	-0.024
Romania	0.262	-0.003	44.3	4003	99.0	0.033	-0.010
Estonia	0.185	-0.003	56.6	3757	513.6	0.102	-0.013
Portugal	0.408	-0.002	60.9	3996	502.5	0.004	-0.002
Ireland	0.486	-0.002	58.8	3648	882.8	0.013	-0.004
France	0.497	-0.002	68.7	3899	561.6	0.006	-0.002
Sweden	0.275	-0.002	67.3	3691	1191.1	0.008	-0.003
South Korea	0.431	-0.002	92.6	5193	386.2	0.004	-0.002
Slovakia	0.151	-0.002	0.0	3840	193.7	-0.049	0.020
Slovenia	0.130	-0.002	94.9	3848	NA	0.012	-0.005
Japan	0.402	-0.001	88.4	5823	353.8	0.000	-0.000
USA	0.372	-0.001	74.5	5620	124.6	0.001	-0.000
Canada	0.448	-0.000	0.0	3958	167.5	0.001	-0.000
Luxembourg	0.242	-0.000	0.0	3691	642.4	-0.011	0.002
China	0.284	-0.000	0.0	5478	322.1	-0.002	0.001
India	0.107	-0.000	0.0	5629	229.4	-0.002	0.000
Turkey	0.245	-0.000	0.0	4341	38.5	-0.008	0.004
Indonesia	0.061	-0.000	0.0	7104	155.4	-0.003	0.001
Mexico	0.319	-0.000	0.0	4963	54.1	-0.001	0.000
Brazil	0.134	-0.000	0.0	6608	130.1	-0.001	0.000
Australia	0.127	-0.000	0.0	8340	69.7	-0.002	0.001
Weighted Average	0.411	-0.003				0.009	-0.003

Notes. This table reports the impact of the closure of Russian and Ukrainian airspace. κ_o^{2019} is the connectedness of country o with its partners before the closure and $\Delta\kappa_o$ its change, both averaged across partners weighted by their GDP. % RUS is the share of that decline attributable to pairs with Russia, the remainder coming from detours between third countries; it is zero for Russia itself, which cannot be its own partner, for the countries that did not ban Russian carriers, and for the banning countries that had no direct-flight connectedness with Russia in 2019, whose declines come entirely from detours; d_{gc} is the GDP-weighted great-circle distance to partners in miles, and $\Delta d^{\text{affected}}$ the average increase in flight distance over the routes that must reroute, which is undefined (infinite) for Russia, whose affected routes are severed rather than lengthened, and unavailable for countries with no rerouted route. It is not available for Cyprus and Slovenia, because the population weights of their airport grids are not available. Australia and Brazil had no direct-flight connectedness with Russia in 2019 and are affected only through detours; their changes round to zero. s_{dom} is the domestic expenditure share in total absorption and w/P the real wage, in which the price index is calculated as $P_d = \prod_s P_{ds}^{\xi_{ds}}$. Countries are sorted by $\Delta\kappa_o$. The last row reports the weighted average over the countries shown, using the base-period GDP as the weights.

Finally, we examine how Russia's trade is reallocated. Because communication costs rise with all of Russia's banning partners at once but not with the partners that remain connected, the model predicts both reshoring and a reallocation of the remaining foreign trade toward those partners. Both are present. Restricting attention to tradable goods (WIOD sectors 1–16), our counterfactual predicts that Russia's domestic share of goods absorption rises from 66.7% to 67.2%, and the domestic share of its sales from 76.6% to 76.8% (the first row of Online Appendix Table D-3).

Figure 5 aggregates partners into groups. Among the countries in the WIOD sample, those that remain connected to Russia are China, India, Turkey and Indonesia, which form the first group. We divide the remaining countries, all subject to the mutual ban, into Western Europe, Eastern Europe and the Baltics, East Asia (Japan and South Korea), North America (the United States and Canada), and the rest of the world, which is the WIOD residual region. Shares are expressed over foreign trade only, which removes the reshoring margin from the denominator and isolates the reallocation across import sources and export destinations: the still-connected group gains 0.56 percentage points of Russia's foreign imports, Western Europe loses on both sides (0.81 percentage points of imports and 1.41 of exports), and the rest of the world absorbs 1.60 percentage points more of Russia's exports. For each group, both margins move largely in the same direction, and the reallocation is typically more pronounced on the export side than on the import side.

Figure 5: Reallocation of Russia's goods trade by partner group



Notes. Goods trade only (WIOD sectors 1–16). The figure reports the change in the share of Russia's goods imports and exports accounted for by each group of partners, in percentage points. Shares are computed over foreign partners only, excluding Russia's domestic flows, so that the figure describes the reallocation of the foreign trade basket rather than the reshoring margin. Online Appendix Table D-3 reports the underlying numbers. The groups are: still-connected partners, i.e. those that did not ban Russian carriers (China, India, Turkey, Indonesia); Western Europe (Germany, France, Italy, the Netherlands, Belgium, the United Kingdom, Spain, Portugal, Ireland, Luxembourg, Austria, Denmark, Finland, Sweden, Greece, Malta, Cyprus); Eastern Europe and the Baltics (the Czech Republic, Poland, Hungary, Romania, Bulgaria, Slovenia, Slovakia, Estonia, Latvia, Lithuania); East Asia (Japan, Korea); North America (the United States, Canada); and the rest of the world, which is the WIOD residual region. All groups other than the first are subject to the mutual ban.

7 Conclusions

This paper demonstrates that communication costs have an impact on international trade. The increase in direct flights reduces the cost of face-to-face communication and product failures, and increases the volume of bilateral trade. We provide both a theoretical framework and empirical evidence to support our argument. We exploit exogenous variation caused by regulations and technology and use an instrumental variable approach to overcome identification challenges. The positive effects of direct flights are larger for those products with a lower elasticity of substitution, further supporting the mechanism by which direct flights facilitate imports by reducing communication costs, as the benefits of fewer product failures are larger for products that are harder to substitute. We also provide direct evidence that air connectedness reduces the probability of product failure.

Embedding our model in general equilibrium, we conduct a counterfactual analysis to calculate the welfare gain from improving international direct flight connectedness. The same framework evaluates shocks that run in the opposite direction. Applying it to the closure of Russian and Ukrainian airspace after February 2022, we find that Russia loses 0.14% of its real wage, reshores part of its goods absorption, and reallocates its remaining foreign trade toward the partners that did not ban its carriers, while the countries that

imposed the ban lose an order of magnitude less. Our approach can be applied to other episodes in which airspace changes.

Declaration of generative AI usage

During the preparation of this work the authors used generative AI tools (Claude Code) to assist with coding and with editing the manuscript. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the article.

References

Allen, Treb and Costas Arkolakis, "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, August 2014, *129* (3), 1085–1140.

Anderson, James E. and Eric Van Wincoop, "Trade Costs," *Journal of Economic literature*, 2004, *42* (3), 691–751.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, 2012, *102* (1), 94–130.

Basker, Emek and Fariha Kamal, "Recall and Response: Relationship Adjustments to Adverse Information Shocks," *European Economic Review*, October 2021, *139*, 103903.

Bernard, Andrew B., Andreas Moxnes, and Yukiko U. Saito, "Production Networks, Geography, and Firm Performance," *Journal of Political Economy*, September 2018, *127* (2), 639–688.

Besedeš, Tibor, Jing Chu, and Antu Panini Murshid, "Fly the Unfriendly Skies: The Role of Transport Costs in Gravity Models of Trade," *Journal of International Economics*, November 2024, *152*, 103994.

Blonigen, Bruce A. and Anca D. Cristea, "Air Service and Urban Growth: Evidence from a Quasi-Natural Policy Experiment," *Journal of Urban Economics*, 2015, *86*, 128–146.

Boehm, Johannes, "The Impact of Contract Enforcement Costs on Value Chains and Aggregate Productivity," *The Review of Economics and Statistics*, January 2022, *104* (1), 34–50.

Broda, Christian and David E. Weinstein, "Globalization and the Gains From Variety," *The Quarterly Journal of Economics*, May 2006, *121* (2), 541–585.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 2015, *82* (1), 1–44.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Campante, Filipe and David Yanagizawa-Drott, "Long-Range Growth: Economic Development in the Global Network of Air Links," *The Quarterly Journal of Economics*, August 2018, *133* (3), 1395–1458.

Chen, Maggie Xiaoyang and Chuanhao Lin, "Geographic Connectivity and Cross-Border Investment: The Belts, Roads and Skies," *Journal of Development Economics*, September 2020, *146*, 102469.

Coşar, A. Kerem and Banu Demir, "Domestic Road Infrastructure and International Trade: Evidence from Turkey," *Journal of Development Economics*, January 2016, *118*, 232–244.

Costinot, Arnaud and Andrés Rodríguez-Clare, "Chapter 4 - Trade Theory with Numbers: Quantifying the Consequences of Globalization," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Handbook of International Economics*, Vol. 4 of *Handbook of International Economics*, Elsevier, January 2014, pp. 197–261.

Cristea, Anca D., "Buyer-Seller Relationships in International Trade: Evidence from U.s. States' Exports and Business-Class Travel," *Journal of International Economics*, July 2011, *84* (2), 207–220.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Unbalanced Trade," NBER Working Paper w13035 April 2007.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, April 2018, *108* (4-5), 899–934.

__ **and Richard Hornbeck**, "Railroads and American Economic Growth: A "Market Access" Approach," *The Quarterly Journal of Economics*, 2016, *131* (2), 799–858.

Ducruet, César, Réka Juhász, Dávid Krisztián Nagy, and Claudia Steinwender, "All Aboard: The Effects of Port Development," *Journal of International Economics*, 2024, *151*, 103963.

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

Fan, Jingting, Yi Lu, and Wenlan Luo, "Valuing Domestic Transport Infrastructure: A View from the Route Choice of Exporters," *Review of Economics and Statistics*, 2023, *105* (6), 1562–1579.

Fieler, Ana Cecília and Jonathan Eaton, "The Margins of Trade," *Econometrica*, 2025, *93* (1), 129–160.

Furusawa, Taiji, Tomohiko Inui, Keiko Ito, and Heiwai Tang, "Global Sourcing and Domestic Production Networks," SSRN Scholarly Paper ID 3043640, Social Science Research Network, Rochester, NY September 2017.

Gates, Dominic, "Boeing 787's Problems Blamed on Outsourcing, Lack of Oversight," *The Seattle Times*, February 2013.

Giri, Rahul, Kei-Mu Yi, and Hakan Yilmazkuday, "Gains from Trade: Does Sectoral Heterogeneity Matter?," Technical Report w26741, National Bureau of Economic Research February 2020.

Grosz, Michel and Devesh Raval, "Fraud Across Borders," Working Paper 2022.

Guillaume, Gaulier and Soledad Zignago, "BACI: International Trade Database at the Product-Level. The 1994-2007 Version," Working Papers 2010-23, CEPII Research Center 2010.

Gumpert, Anna, "The Organization of Knowledge in Multinational Firms," *Journal of the European Economic Association*, December 2018, *16* (6), 1929–1976.

Harrigan, James, "OECD Imports and Trade Barriers in 1983," *Journal of International Economics*, August 1993, *35* (1–2), 91–111.

Hart, Peter E., Nils J. Nilsson, and Bertram Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE transactions on Systems Science and Cybernetics*, 1968, *4* (2), 100–107.

Hummels, David, "Have International Transportation Costs Declined?," Working Paper 1999.

Imbens, Guido W. and Thomas Lemieux, "Regression Discontinuity Designs: A Guide to Practice," *Journal of econometrics*, 2008, *142* (2), 615–635.

Liker, Jeffrey and Thomas Y. Choi, "Building Deep Supplier Relationships," *Harvard Business Review*, December 2004.

Mayer, Thierry and Soledad Zignago, "Notes on CEPII's Distances Measures: The GeoDist Database," SSRN Scholarly Paper 1994531 December 2011.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, February 2008, *142* (2), 698–714.

Nunn, Nathan, "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade," *The Quarterly Journal of Economics*, January 2007, *122* (2), 569–600.

Poole, Jennifer Pamela, "Business Travel as an Input to International Trade," Technical Report 2009.

Rauch, James E., "Networks Versus Markets in International Trade," *Journal of International Economics*, June 1999, *48* (1), 7–35.

Simonovska, Ina and Michael E. Waugh, "The Elasticity of Trade: Estimates and Evidence," *Journal of International Economics*, January 2014, *92* (1), 34–50.

Soderbery, Anson, "Estimating Import Supply and Demand Elasticities: Analysis and Implications," *Journal of International Economics*, May 2015, *96* (1), 1–17.

Söderlund, Bengt, "The Importance of Business Travel for Trade: Evidence from the Liberalization of the Soviet Airspace," *Journal of International Economics*, November 2023, *145*, 103812.

Startz, Meredith, "The Value of Face-to-Face: Search and Contracting Problems in Nigerian Trade," Working Paper 2025.

Tang, Christopher S., Joshua D. Zimmerman, and James I. Nelson, "Managing New Product Development and Supply Chain Risks: The Boeing 787 Case," *Supply Chain Forum: An International Journal*, January 2009, *10* (2), 74–86.

Timmer, Marcel P., Erik Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J. de Vries, "An Illustrated User Guide to the World Input–Output Database: The Case of Global Automotive Production," *Review of International Economics*, 2015, *23* (3), 575–605.

Wang, Linfeng and Shihe Fu, "Fly to Trade: Effects of International Direct Flights on Chinese Cities' Trade," SSRN Scholarly Paper 4191988, Rochester, NY August 2022.

Yilmazkuday, Demet and Hakan Yilmazkuday, "The Role of Direct Flights in Trade Costs," *Review of World Economics*, May 2017, *153* (2), 249–270.

Online Appendix

A Additional Theoretical Results

Table A-1 collects the notation used in Section 2.

Table A-1: Notation

Symbol	Description
$o, d \in \mathcal{N}$	Origin and destination countries
$s, k \in \mathcal{S}$	Sectors
$\omega \in [0,1]$	Intermediate goods variety
Q_s, $x(\omega)$	Sectoral output of the final goods producer; quantity of variety ω (equation (1))
ρ_s	Elasticity of substitution among input varieties
$\tilde{\rho}_s \equiv \frac{\rho_s}{\rho_s - 1}$	Degree of product differentiation
$z_{os}(\omega)$	Productivity draw of the supplier of variety ω in country o, sector s
T_{os}, θ_s	Location and dispersion parameters of the Fréchet productivity distribution
c_{os}	Unit cost of the input bundle in country o, sector s (equation (12))
t_{od}	Iceberg trade cost
α, α_{ods}	Probability that an input meets the buyer's standard (communication intensity); its optimal common value for inputs sourced from o
m_{od}	Exogenous communication cost; raising α costs $e^{m_{od}\alpha}$ per unit
$G_{ods}(p)$	Distribution of supplier prices induced by the productivity draws
P_{ods}	Price index of the composite input sourced solely from country o (equation (3))
P_{ds}, γ_s	Price index of the sector-s composite input bundle in country d; its constant (equation (6))
X_{ods}, X_{ds}	Trade flow from o to d in sector s; total absorption in d, s
δ_{os}, δ_{ds}	Origin-sector and destination-sector fixed effects (equation (7))

A.1 Alternative Ways the Reliability Margin Enters Demand

In our model the success probability α_{ods} is a probability and therefore enters the production function linearly, multiplying the input aggregator in the buyer's problem (2). A natural question is whether this matters. Mathematically, it acts like a demand shifter, and one may interpret it as "quality", as in Fieler and Eaton (2025) and Broda and Weinstein (2006). We show here that the way α enters demand determines the elasticity of trade with respect to communication, and that some natural alternatives are inconsistent with our finding that communication frictions reduce trade more in sectors with a lower ρ_s (Section 5).

Suppose the buyer's effective input from variety ω is $\alpha(\omega)^{\eta_s} x(\omega)$ for some exponent $\eta_s > 0$,

so that the sectoral aggregator is

$$Q_s = \left(\int_0^1 \left(\alpha(\omega)^{\eta_s} x(\omega) \right)^{(\rho_s - 1)/\rho_s} d\omega \right)^{\rho_s/(\rho_s - 1)}, \tag{A-1}$$

while, exactly as in the baseline, raising $\alpha(\omega)$ costs $e^{m_{od}\alpha(\omega)}$ per unit, so that $p_{ods}(\omega) = e^{m_{od}\alpha(\omega)} c_{os} t_{od} / z_{os}(\omega)$.[27] Our baseline is the special case $\eta_s = \tilde{\rho}_s = \rho_s/(\rho_s - 1)$. Fieler and Eaton (2025) consider a quality that is substitutable with quantity, so $\eta_s = 1$ and consumer cares about the composite $\alpha(\omega)x(\omega)$. Broda and Weinstein (2006) and many other papers set $\eta_s = 1/(\rho_s - 1)$, so the demand shifter $\alpha(\omega)$ enters the price index linearly.

The exponent η_s has an economic interpretation in terms of what fails. In our model a variety either meets the buyer's standard or is useless: with a continuum of varieties, a fraction $1 - \alpha$ of the varieties sourced from o fails entirely, the surviving measure is α, and the aggregator integrates over the surviving varieties only,

$$Q_s = \left(\int_0^1 \alpha(\omega)\, x(\omega)^{(\rho_s - 1)/\rho_s} d\omega \right)^{\rho_s/(\rho_s - 1)},$$

which is (A-1) with $\eta_s = \tilde{\rho}_s$, since $\alpha^{\tilde{\rho}_s(\rho_s - 1)/\rho_s} = \alpha$. Suppose instead that a fraction $1 - \alpha$ of the units ordered of a variety fails, so that the buyer's effective quantity is $\alpha(\omega)\, x(\omega)$ and the aggregator is

$$Q_s = \left(\int_0^1 \left(\alpha(\omega)\, x(\omega) \right)^{(\rho_s - 1)/\rho_s} d\omega \right)^{\rho_s/(\rho_s - 1)},$$

which is (A-1) with $\eta_s = 1$. This is the substitutable-quality formulation of Fieler and Eaton (2025) once their quality index is read as a success probability in $(0, 1]$. The two formulations therefore differ in what a failure destroys: a whole variety in our model, a fraction of the units of a variety in theirs.

Cost minimization over $\{x(\omega)\}$ and then over $\{\alpha(\omega)\}$ proceeds exactly as in Section 2. The quality index enters the unit-cost problem as the weight $\alpha(\omega)^{\eta_s}$, so the buyer chooses α to minimize $\alpha^{-\eta_s} e^{m_{od}\alpha}$, yielding the interior optimum

$$\alpha_{ods} = \frac{\eta_s}{m_{od}}, \tag{A-2}$$

common across suppliers and independent of the productivity draw. The effective unit cost of sourcing ω from o is $(e\, m_{od}/\eta_s)^{\eta_s}\, c_{os} t_{od} / z_{os}(\omega)$, and the Fréchet algebra delivers the gravity equation

$$\log X_{ods} = \delta_{ds} + \log T_{os} - \theta_s \log c_{os} - \theta_s\, \eta_s\, \log m_{od} - \theta_s \log t_{od}. \tag{A-3}$$

[27] In the appendix, we deviate from the notation in the paper slightly, and write $\alpha(\omega)$ instead of $\alpha(p(\omega))$.

The elasticity of trade with respect to the communication friction is thus $\theta_s \eta_s$.

In our case, $\eta_s = \tilde{\rho}_s$: failure by the variety delivers $\alpha_{ods} = \tilde{\rho}_s / m_{od}$ and the elasticity $\theta_s \tilde{\rho}_s$, so industries with lower ρ_s will be affected more by the communication costs, because losing a variety is costlier the more differentiated the inputs are. When we interpret $\alpha(\omega)$ as the substitutable quality in Fieler and Eaton (2025), $\eta_s = 1$: failure by the unit delivers $\alpha_{ods} = 1/m_{od}$ and the elasticity θ_s, so the communication friction acts as an iceberg cost and ρ_s does not enter the gravity equation at all. Finally, if we follow the convention in Broda and Weinstein (2006) and set $\eta_s = 1/(\rho_s - 1)$, we again have industries with lower ρ_s affected more by the communication costs. We prefer our setup because it is consistent with the gravity results and with the nature of the product failures in the lawsuit data, which involve the whole product rather than a fraction of the units delivered (Online Appendix B.3).

A.2 General Communication Costs

In the baseline model the communication cost takes the convenient exponential form $e^{m_{od}\alpha(\omega)}$. Here we show that the central comparative static—that communication frictions reduce trade more in sectors with a lower elasticity of substitution—holds for any communication cost that is log-supermodular in m_{od} and $\alpha(\omega)$. Log-supermodularity means that the percentage change in the communication cost in response to a higher $\alpha(\omega)$ is increasing in m_{od}: a worse communication environment makes additional reliability proportionally more expensive. Because buyers in low-substitution sectors optimally choose higher reliability, they are the most exposed to a worse communication environment, and their trade falls the most when m_{od} rises.

Replace the exponential cost in the baseline model with a general positive, twice continuously differentiable function $f(m_{od}, \alpha) > 0$, where m_{od} is the exogenous communication friction and α is the probability that an input meets the buyer's standard, so that the price of variety ω sourced from o is $p_{ods}(\omega) = f(m_{od}, \alpha(\omega))\, c_{os} t_{od} / z_{os}(\omega)$. Repeating the cost minimization that delivers (3), conditional on sourcing from o the buyer chooses α to minimize $\alpha^{-\tilde{\rho}_s} f(m_{od}, \alpha)$, where $\tilde{\rho}_s \equiv \rho_s/(\rho_s - 1)$ as in the main text. Define the minimized communication wedge

$$\mathcal{C}(\tilde{\rho}_s, m_{od}) \equiv \min_{\alpha \in (0,1]} \alpha^{-\tilde{\rho}_s} f(m_{od}, \alpha), \tag{A-4}$$

and, writing $g(m, \alpha) \equiv \log f(m, \alpha)$, its logarithm

$$V(\tilde{\rho}_s, m_{od}) \equiv \log \mathcal{C}(\tilde{\rho}_s, m_{od}) = \min_{\alpha \in (0,1]} \left\{ g(m_{od}, \alpha) - \tilde{\rho}_s \log \alpha \right\}. \tag{A-5}$$

The wedge $\mathcal{C}$ plays the role of the factor $(e\, m_{od}/\tilde{\rho}_s)^{\tilde{\rho}_s}$ in the baseline model: the effective

unit cost of sourcing ω from o is $\mathcal{C}(\tilde{\rho}_s, m_{od})\, c_{os} t_{od} / z_{os}(\omega)$.

We impose the following regularity condition.

Assumption 1 *The function $g(m, \alpha)$ is twice continuously differentiable, and for every relevant pair $(\tilde{\rho}_s, m_{od})$ the problem in (A-5) has a unique interior solution $\alpha_{ods} \equiv \alpha^*(\tilde{\rho}_s, m_{od}) \in (0, 1)$ satisfying the strict second-order condition*

$$g_{\alpha\alpha}(m_{od}, \alpha_{ods}) + \frac{\tilde{\rho}_s}{\alpha_{ods}^2} > 0. \tag{A-6}$$

The first-order condition for the optimal success probability is

$$g_\alpha(m_{od}, \alpha_{ods}) = \frac{\tilde{\rho}_s}{\alpha_{ods}}, \qquad \text{equivalently} \qquad \alpha_{ods}\, g_\alpha(m_{od}, \alpha_{ods}) = \tilde{\rho}_s. \tag{A-7}$$

Totally differentiating (A-7) and using (A-6),

$$\frac{\partial \alpha_{ods}}{\partial \tilde{\rho}_s} = \frac{1}{\alpha_{ods}\left[g_{\alpha\alpha}(m_{od}, \alpha_{ods}) + \tilde{\rho}_s / \alpha_{ods}^2\right]} > 0, \tag{A-8}$$

so sectors with a larger $\tilde{\rho}_s$ (a lower elasticity of substitution ρ_s) choose a higher success probability, exactly as in the baseline model where $\alpha_{ods} = \tilde{\rho}_s / m_{od}$.

Proposition 1 *Under Assumption 1,*

$$\frac{\partial^2 V(\tilde{\rho}_s, m_{od})}{\partial \tilde{\rho}_s\, \partial \log m_{od}} = \frac{m_{od}\, g_{m\alpha}(m_{od}, \alpha_{ods})}{\alpha_{ods}\left[g_{\alpha\alpha}(m_{od}, \alpha_{ods}) + \tilde{\rho}_s / \alpha_{ods}^2\right]}, \tag{A-9}$$

and therefore, evaluated at $\alpha = \alpha_{ods}$,

$$\frac{\partial^2 V(\tilde{\rho}_s, m_{od})}{\partial \tilde{\rho}_s\, \partial \log m_{od}} \gtrless 0 \quad \Longleftrightarrow \quad \frac{\partial^2 \log f(m_{od}, \alpha)}{\partial m_{od}\, \partial \alpha} \gtrless 0. \tag{A-10}$$

Proof. By the envelope theorem applied to (A-5), $\partial V / \partial \log m_{od} = m_{od}\, g_m(m_{od}, \alpha_{ods})$. Differentiating with respect to $\tilde{\rho}_s$ and substituting (A-8),

$$\frac{\partial^2 V}{\partial \tilde{\rho}_s\, \partial \log m_{od}} = m_{od}\, g_{m\alpha}(m_{od}, \alpha_{ods})\, \frac{\partial \alpha_{ods}}{\partial \tilde{\rho}_s} = \frac{m_{od}\, g_{m\alpha}(m_{od}, \alpha_{ods})}{\alpha_{ods}\left[g_{\alpha\alpha}(m_{od}, \alpha_{ods}) + \tilde{\rho}_s / \alpha_{ods}^2\right]}.$$

Because $m_{od} > 0$, $\alpha_{ods} > 0$, and the bracketed second-order term is positive by (A-6), the sign of the cross-partial equals the sign of $g_{m\alpha} = \partial^2 \log f / \partial m\, \partial \alpha$. □

Proposition 1 identifies the relevant primitive as the *log-supermodularity* of the commu-

nication cost function,

$$\frac{\partial^2 \log f(m_{od}, \alpha)}{\partial m_{od}\, \partial \alpha} > 0, \tag{A-11}$$

which states that a higher communication friction raises the marginal (log) cost of achieving a higher success probability. The baseline cost $f = e^{m_{od}\alpha}$ satisfies (A-11) with $\partial^2 \log f / \partial m_{od}\, \partial \alpha = 1$.

To translate this into trade flows, the generalized gravity equation replaces the term $\theta_s \tilde{\rho}_s \log m_{od}$ in (7) with $\theta_s V(\tilde{\rho}_s, m_{od})$:

$$\log X_{ods} = \delta_{ds} + \log T_{os} - \theta_s \log c_{os} - \theta_s \log t_{od} - \theta_s V(\tilde{\rho}_s, m_{od}), \tag{A-12}$$

where, as in (7), the destination-sector fixed effect δ_{ds} absorbs the price index and any origin-invariant constants. The elasticity of bilateral trade with respect to the communication friction is

$$\frac{\partial \log X_{ods}}{\partial \log m_{od}} = -\theta_s\, m_{od}\, g_m(m_{od}, \alpha_{ods}) < 0, \tag{A-13}$$

and, by Proposition 1,

$$\frac{\partial^2 \log X_{ods}}{\partial \tilde{\rho}_s\, \partial \log m_{od}} = -\theta_s \frac{m_{od}\, g_{m\alpha}(m_{od}, \alpha_{ods})}{\alpha_{ods}\left[g_{\alpha\alpha}(m_{od}, \alpha_{ods}) + \tilde{\rho}_s/\alpha_{ods}^2\right]}. \tag{A-14}$$

Since $\tilde{\rho}_s = \rho_s/(\rho_s - 1)$ is decreasing in ρ_s, under log-supermodularity (A-11) we have

$$\frac{\partial^2 \log X_{ods}}{\partial \rho_s\, \partial \log m_{od}} = \theta_s \frac{m_{od}\, g_{m\alpha}(m_{od}, \alpha_{ods})}{(\rho_s - 1)^2\, \alpha_{ods}\left[g_{\alpha\alpha}(m_{od}, \alpha_{ods}) + \tilde{\rho}_s/\alpha_{ods}^2\right]} > 0. \tag{A-15}$$

Hence the trade elasticity with respect to communication costs becomes less negative as ρ_s rises: communication frictions reduce trade more strongly in sectors with a lower elasticity of substitution.

It is the interaction between m_{od} and α in the cost function that drives this result. If the communication cost is multiplicatively separable, $f(m_{od}, \alpha) = A(m_{od})B(\alpha)$, then $g_{m\alpha} = 0$ and

$$\mathcal{C}(\tilde{\rho}_s, m_{od}) = A(m_{od}) \min_{\alpha \in (0,1]} \alpha^{-\tilde{\rho}_s} B(\alpha),$$

so the effect of m_{od} on trade is the same in every sector: the bilateral friction and the sectoral elasticity do not interact. The exponential baseline is not separable in this sense, which is precisely why it generates the $\tilde{\rho}_s$-amplification.

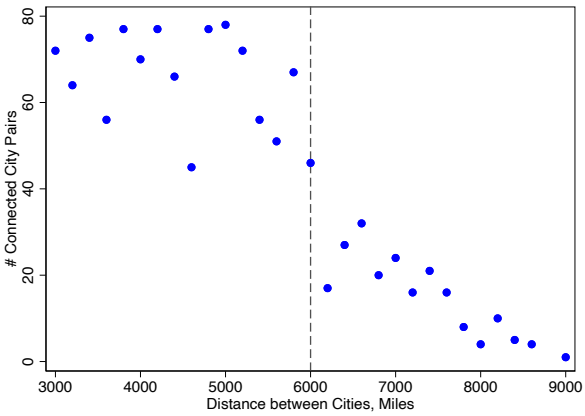
B Additional Empirical Results

B.1 Discontinuity of Connectedness at 6,000 Miles

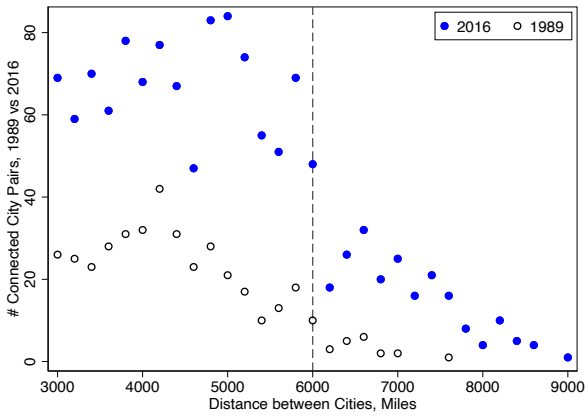
Following Campante and Yanagizawa-Drott (2018), we find that air connectedness between cities drops discontinuously around 6,000 miles due to changes in policy and aviation technology. The United States and Europe passed regulations in the 1950s and 1991, respectively, that restrict hours of operation for flight crew members within a 24-hour period. The regulations require an additional crew member on any flight exceeding 12 hours. Given the speeds of civil aviation aircraft, a 12-hour flight translates into about 6,000 miles. Therefore, we speculate that there will be far fewer direct flights over 6,000 miles than within 6,000 miles. On the other hand, before the Boeing 747-400 aircraft was introduced to the market in 1989, there were very few aircraft that could afford long-haul flights of 9-12 hours. The arrival of the long-haul aircraft makes the 12-hour flight limit more meaningful: the discontinuity at 6,000 miles should be exacerbated by aircraft capable of long-distance flying.

Figure B-1 confirms the existence of the discontinuity above. Panel B-1a shows a sharp drop in the number of connected city pairs when the distance is larger than 6,000 miles in 2016. Panel B-1b also adds the dots to represent the number of connected city pairs in 1989. We can see the drop is much smoother back in 1989. And in the range of 4,600 to 6,000 miles, the number of connected city pairs dropped significantly more in 1989 than in 2016. Panel B-1c shows the potential number of connected city pairs across country: there is no sharp drop at 6,000 miles.

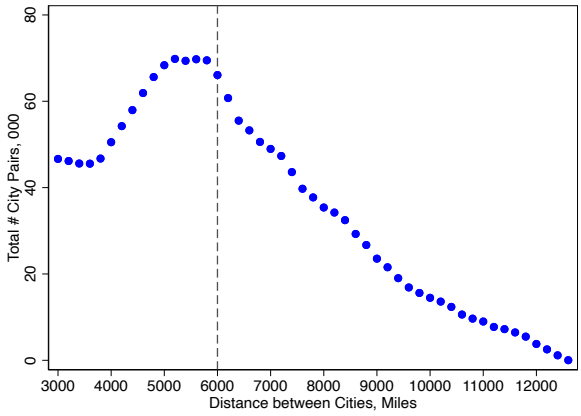
Figure B-1: Connections between City Pairs, by Distance



(a) 2016



(b) 1989 vs 2016



(c) Potential Connected City Pairs

Notes. We group grid pairs into 200-mile bins. Each dot represents the number of connected cities in the preceding 200 miles. City pairs in panel B-1c represent the permutations of the 2,197 cities around the world that have airports in the TFS data and for which exact coordinates can be found. We have 4,640,548 potential connected city pairs after excluding all the domestic pairs.

B.2 Summary Statistics

Table B-1: List of Country-pairs with the Highest Number of Cases, Defective Product

Rank	Country Pair	# Litigation Cases	Rank	Country Pair	# Litigation Cases
1	DEU-USA	31	16	BEL-USA	9
2	GBR-USA	24	17	USA-FRA	8
3	FRA-USA	21	18	USA-JPN	7
4	USA-DEU	20	19	IND-USA	6
5	ITA-USA	20	20	USA-TWN	5
6	AUS-USA	19	21	USA-SWE	5
7	NLD-USA	19	22	NOR-USA	4
8	CAN-USA	15	23	CHN-USA	4
9	USA-GBR	13	24	AUT-USA	3
10	USA-ITA	13	25	POL-USA	3
11	USA-CAN	13	26	NLD-ITA	3
12	USA-NLD	11	27	DEU-ITA	3
13	USA-AUS	11	28	NZL-DEU	2
14	JPN-USA	10	29	USA-BEL	2
15	SWE-USA	9	30	RUS-USA	2

Notes. The Lexis US data includes 437 country pairs with some "defective product" lawsuits. Here, we list only the 30 country pairs with the highest number of such lawsuits. The first country refers to the location of the plaintiffs, while the second country refers to the location of the defendants. The average number of cases between countries is 1 when searching for "defective product". In 339 country pairs, both the plaintiffs and the defendants are located outside the US.

Table B-2: List of Country-pairs with the Highest Number of Foreign Ownership Links

Rank	Country of GUO	Country of Subsidiary	# Foreign Ownership Links
1	United States	United Kingdom	15186
2	United States	France	5247
3	United Kingdom	United States	4165
4	United States	Germany	4084
5	Netherlands	United Kingdom	3880
6	United States	Canada	3826
7	Belgium	France	3807
8	Japan	United States	3678
9	Netherlands	Belgium	3444
10	Netherlands	Germany	3315
11	Netherlands	France	3259
12	Germany	France	3182
13	France	United Kingdom	3134
14	Canada	United States	2863
15	Germany	United Kingdom	2846
16	France	Belgium	2795
17	United Kingdom	France	2611
18	United States	Italy	2540
19	Italy	France	2504
20	Germany	United States	2485
21	Netherlands	Spain	2457
22	Germany	Italy	2443
23	France	Italy	2403
24	Germany	Austria	2343
25	Germany	Czechia	2145
26	United Kingdom	Germany	2121
27	Germany	Poland	2105
28	Austria	Germany	2068
29	United States	Spain	2043
30	France	United States	2008

Notes. From 2011 to 2017, we identify a total of 270,096 foreign ownership links distributed across 2,654 country pairs in ORBIS. On average, there are 102 foreign ownership links between two countries with a positive number of links. We list 30 country pairs that have the largest number of foreign ownership links.

B.3 Textual and Case Analysis of Defective-Product Cases

This section examines the content of the "defective product" case documents (we omit the quotation marks hereafter). We first describe how we construct the firm-vs-firm sample from the raw documents. We then provide two complementary pieces of analysis: a keyword analysis of the full sample, which confirms that the cases are disputes over failed manufactured products rather than services; and a detailed reading of representative cases, which shows that the disputes typically center on product specifications and use conditions that were not fully communicated between buyer and seller.

To construct the sample, instead of extracting only the two parties named in the case title, we download the full case documents and extract all plaintiffs and defendants from

each document's header. We keep only firm-vs-firm cases, in which at least one plaintiff and at least one defendant carry a clear company suffix such as Inc., Co., or Ltd., so that the sample involves upstream suppliers and downstream firms.[28] When matching party names to firms in ORBIS, we construct a comprehensive name-cleaning process based on the nature of firm names in both datasets. The matching rate for firms involved in defective-product cases is 77%.[29] The resulting sample contains 559 cross-country defective-product cases (with an effective number of 384) matched to ORBIS, distributed across 437 country pairs.

Because the Lexis records typically contain only abbreviated firm names, a single Lexis party may match multiple ORBIS firms that share the same cleaned name but are distinct entities located in different countries. Each ORBIS firm is mapped to a unique global ultimate owner (GUO) through the ORBIS controlling-shareholder hierarchy, so a one-to-many Lexis–ORBIS match implies that one Lexis party can be associated with multiple ultimate owners located in different grids. We handle this multiplicity through a weighting scheme: if the plaintiff in a given case maps to n_1 ultimate-owner grids and the defendant maps to n_2 ultimate-owner grids, the case generates $n_1 \times n_2$ plaintiff–defendant grid-pair observations, each assigned an allocation weight of $1/(n_1 \times n_2)$, so that the allocation weights of a given case sum to one. Each grid-pair observation then carries the product of this allocation weight and the case's firm-vs-firm weight, the share of plaintiff–defendant pairs that are firm pairs (footnote 28); the weights of a given case thus sum to its effective number. Summing these weights across cases within each grid pair yields the effective number of cases between two grids, which is the dependent variable in our grid-pair level regressions.

The keyword analysis confirms that the sample consists of disputes over failed manufactured products. We apply Term Frequency-Inverse Document Frequency (TF-IDF) to the body text of each case document.[30] The TF-IDF score of a term t in document d is defined as $\text{TF-IDF}(t, d) = \text{TF}(t, d) \times \text{IDF}(t)$, where the term frequency $\text{TF}(t, d)$ measures how often t appears in d, and the inverse document frequency $\text{IDF}(t) = \log(N/\text{DF}(t))$ down-

[28] When a case involves multiple parties, we count it by the share of plaintiff–defendant pairs that are firm pairs. For example, if a case has 3 plaintiffs and 4 defendants, of which 2 plaintiffs and 2 defendants are firms, the case is a firm-vs-firm case with an effective number of $\frac{2\times2}{3\times4} = \frac{1}{3}$.

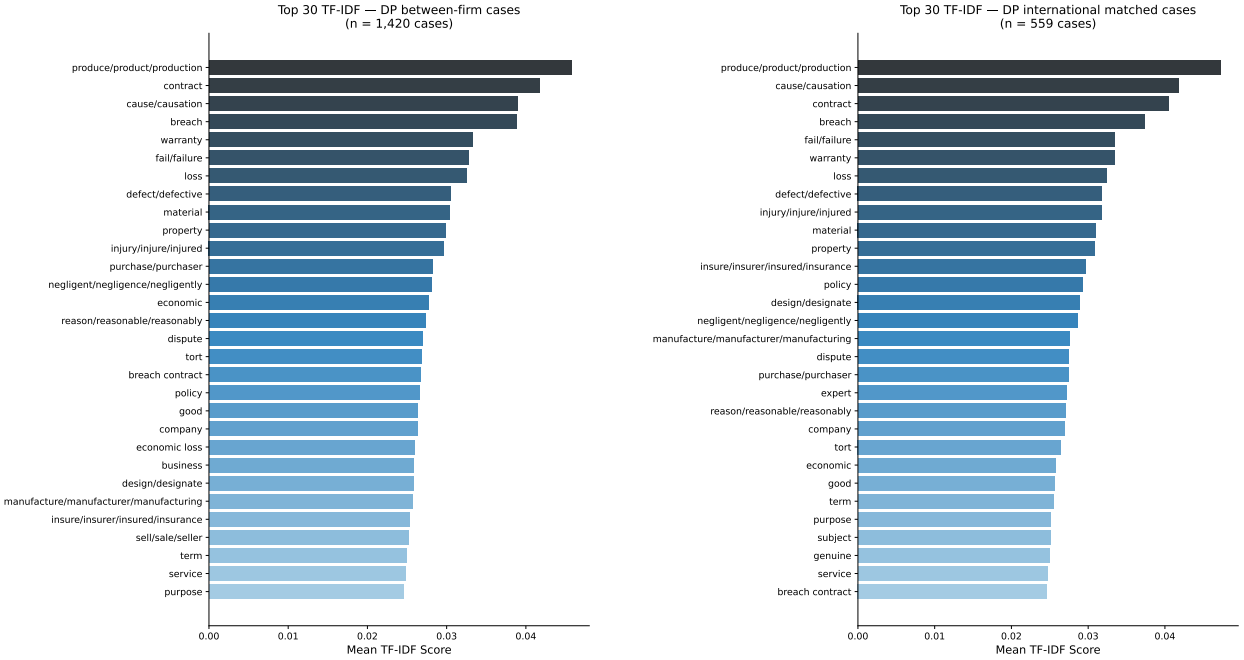
[29] We standardize firm names on both sides by transliterating accented characters to ASCII, converting to uppercase, stripping legal-entity suffixes (e.g., Inc., Corp., Ltd., LLC) using the `cleanco` library, and normalizing common connectors (e.g., "and" to "&"), and we then perform an exact merge on the cleaned names. In addition, we add company suffixes from more countries and languages, handle legal terminology used when listing plaintiffs and defendants, separate alternative company names (appearing as "doing business as", "formerly known as", or in brackets or slashes), and standardize easily confused singular/plural and verb/noun forms, such as construction/constructions/constructing.

[30] We first extract the opinion body from each document, stripping header information (case name, court, counsel listings) and trailing boilerplate (judge signature blocks, dates). The text is then preprocessed by removing numbers, lemmatizing words using their part-of-speech tags, and filtering out English stop words, legal procedural terms (e.g., "deposition", "jurisdiction", "exhibit"), state names, and court citation formats.

weights terms that appear in many documents, such as general legal vocabulary.[31] Figure B-2 presents the top 30 terms ranked by mean TF-IDF score, for all firm-vs-firm cases (panel (a)) and for the subset matched to ORBIS with global ultimate owners in different countries (panel (b)). The ranking is computed on the full set of 1,420 firm-vs-firm cases, not on a selected subset. In both panels, the highest-ranking terms concern product defects (defect/defective, fail/failure), warranties, and product design: nine of the top ten terms coincide across the two panels (product, contract, causation, breach, warranty, failure, loss, defect, material), and the terms that follow (injury, purchase, negligence, design, manufacture, seller, economic loss) point in the same direction. This is the vocabulary of a delivered product that failed, together with that of the causes of action—breach of contract, breach of warranty, negligence—under which firms plead such failures. These keywords are consistent with the cases being disputes over products that failed to meet the buyer's expectations, rather than over services or unrelated matters. The similarity of the two panels indicates that restricting the sample to matched international cases does not systematically alter the nature of the disputes.

[31] Different forms of words, such as verbs, nouns, and adjectives (e.g., defect and defective; fail and failure), are grouped so that different forms of the same word are not counted separately.

Figure B-2: Top-30 Terms by Mean TF-IDF Score in Defective-Product Cases



(a) All firm-vs-firm cases

(b) Matched international cases

Notes. Each bar shows the mean TF-IDF score of a term across all documents in the subsample. TF-IDF is computed using scikit-learn's `TfidfVectorizer` with a minimum document frequency of 5. Text preprocessing includes lemmatization, removal of numbers, English and legal stop words, state names, and court citation formats. Different forms of words, such as verbs, nouns, and adjectives (e.g., defect and defective) are grouped together. Panel (a) includes all 1,420 firm-vs-firm defective-product cases. Panel (b) includes the 559 cases in which both parties are matched to firms in ORBIS with global ultimate owners in different countries.

The hand-coded cases show that the sample covers a wide range of intermediate products and that the litigation typically originates in product specifications. To illustrate, we select 10 cases and code each by reading the full case documents; Table B-3 reports the product involved and the reason for litigation in each case. The products range from industrial machinery and oilfield supplies to precision components and construction materials. In each case, the plaintiff alleges that the delivered product failed to meet specifications or was unsuitable for its intended use.

Table B-3: Selected Defective-Product Cases: Products and Litigation Reasons

Lexis ID	Product	Litigation Reason
LEXIS 200607	Cottonseed processing equipment	Plaintiff alleged the entire cottonseed processing system was defectively manufactured, defectively designed, and not fit for its intended purpose of processing 150 tons per day.

Table B-3: (continued)

Lexis ID	Product	Litigation Reason
LEXIS 204780	Portable oxygen concentrators	Plaintiff sued for breach of contract, breach of express and implied warranty, and fraud, alleging 32 out of 50 units failed to perform as promised and defendant failed to properly repair or replace defective units under the extended 5-year warranty.
LEXIS 190882	Grade P-110 tubular steel coupling stock	Plaintiff sued both distributor and manufacturer for breach of contract, breach of warranties, strict liability, negligence, and negligent misrepresentation, alleging the steel stock was far below the specified P-110 grade, causing well failures.
LEXIS 26039	Electric motors for hydraulic fracturing hot water units	Plaintiff sued for breach of implied warranty of merchantability and fitness for a particular purpose, alleging the motors incorporated into hot water units used for fracking were defective.
LEXIS 800	Crystallizer tanks (26 units)	Plaintiff sued for breach of contract/express warranties, negligent nondisclosure, and negligent misrepresentation, alleging defendant knew but failed to disclose that 304 stainless steel should not be exposed to chlorides, and falsely represented that protective coatings would prevent corrosion.
LEXIS 44610	Valve cores with epichlorohydrin rubber seals for gas cylinders	Plaintiff sued for breach of contract, breach of warranties, and negligent misrepresentation, alleging valve core seals were chemically incompatible with propylene gas, causing leaks and a worldwide product recall.
LEXIS 195115	Polyisobutylene (PIB) window sealant	Plaintiffs alleged the PIB sealant was defectively manufactured, causing it to fail, deteriorate, and migrate within insulating glass units, damaging window components and requiring costly replacements across multiple residential projects.
LEXIS 193964	Ball valves for a natural gas facility	Plaintiff alleged defendant fraudulently modified Class 600 valves to pass them off as Class 900, including scratching nameplates and falsely certifying the valves in pressure test reports.

Table B-3: (continued)

Lexis ID	Product	Litigation Reason
LEXIS 11257	Refractory bricks for industrial glass furnace	Plaintiff sued for breach of contract, breach of warranty, negligence, and product liability, alleging the refractory materials failed to comply with furnace design specifications, causing a furnace wall collapse.
LEXIS 29518	Resin-coated sand (proppant) for well fracture stimulation	Plaintiff alleged defendant supplied and used defective stimulation fluids whose technical specifications were misrepresented, irreparably damaging the well. The manufacturer revised the data sheet one month after the job, confirming the product was used outside its effective range.

Notes. We select 10 defective-product cases from our dataset and hand-code the product involved and the litigation reason by reading the full case documents. Because the case documents are obtained from the proprietary Lexis database, we do not report the parties' names; cases are identified by their Lexis document IDs. The cases illustrate the range of intermediate goods in the sample, including industrial machinery, oilfield supplies, medical facilities, precision components, and construction materials. In each case, the plaintiff alleges that the delivered product failed to meet specifications or was unsuitable for its intended use.

A closer reading of three of these cases illustrates how the disputes connect to the communication friction in our model:

1. LEXIS 800: the seller offered a customization service, designing crystallizer tanks for its clients. The court characterized this process as collaborative in nature, relying on "on-going communications" between both parties rather than one-sided expertise. In practice, however, both sides failed to communicate adequately about material risks. On the seller's side, it did not proactively share its technical knowledge that 304 stainless steel was vulnerable to chloride exposure, or that the protective coating had been applied outside of the manufacturer's specifications. On the buyer's side, it never inquired about chloride risks or material limitations despite actively participating in the collaborative design process. The court found that neither party fulfilled its role in the collaborative process. Critically, the court held that the seller had no duty to disclose the risks it knew about—because the relationship was a standard commercial transaction rather than a fiduciary one, the seller was not obligated to proactively volunteer information about the material's vulnerabilities. The equipment ultimately failed, illustrating how insufficient bilateral communication in a buyer-seller relationship can lead to product defects.

2. LEXIS 44610: the seller designed and supplied valve cores for gas cylinders whose

rubber seal was incompatible with propylene gas. The parties dispute whether this incompatibility was ever definitively communicated to the buyer, which had chosen the seller precisely for its expertise in valve components and relied on the components being suitable for the intended gas types. The incompatibility surfaced only after the product was in use and led to a worldwide product recall.

3. LEXIS 29518: a proppant was selected for a fracture stimulation job on the basis of the producer's published technical data sheet, whose stated closure-stress range covered the anticipated conditions. The data sheet was revised shortly after the job, placing the actual conditions outside the product's effective range. The buyer's engineer had relied solely on the published specifications, and there was no direct communication between the two parties about the product's actual performance limits.

We do not take a stand on which party is responsible in these disputes. What the cases share is that a product ended up in a use case it could not serve because specifications, material limitations, or operating conditions were not fully communicated between buyer and seller. In several cases, a single piece of information—a protective coating, a gas type, a stress range—would plausibly have prevented the failure. Better communication could therefore have addressed many of these issues, which is precisely the friction our model captures.

A natural objection is that these disputes reach the court as claims of breach of contract or breach of warranty, as the keyword ranking and Table B-3 show, and might therefore be read as evidence of incomplete contracts rather than of communication failures. Had the contract specified the feature that later failed, assigning responsibility would have been straightforward and the dispute might not have been litigated. The three cases above show why the contracts were silent: the chloride exposure of the tanks, the gas the valve seal had to withstand, and the closure stress the proppant would face were not written into the contract because they had not been communicated between the parties before the order. An incomplete specification is the form that a communication failure takes by the time it reaches a court, so the prevalence of contract and warranty claims is consistent with the friction in our model.

B.4 Robustness of the impact of air connectedness on trade

B.4.1 Exclude Products Transported by Air

In this section, we change our sample products to exclude the potential impact of air regulations and aircraft technologies on freight rates. Since we do not have data on transportation

modes for trade between all country pairs, we use the US import data by product and mode from the US Census Bureau [32] to calculate the share of air shipping of each product. In our original sample, approximately half of the products have an air transport share of less than 5%. In Table B-4, we focus on products that are not typically shipped by air, i.e., those with an air transport share below 5% or 1% in trade value and re-run our main specifications. The coefficients are positive and significant at conventional levels.

Table B-4: Robustness Checks: Exclude Products Transported by Air, Share = $\frac{\text{Air Value}}{\text{Total Gen Value}}$

	Drop Products with Air Freight Share of above 5%			Drop Products with Air Freight Share of above 1%		
	OLS	IV	First Stage	OLS	IV	First Stage
	(1)	(2)	(3)	(4)	(5)	(6)
Connection Index$_{od} \times \theta_s\tilde{\rho}_s$	0.027^a	0.057^a		0.023^a	0.056^a	
	(0.003)	(0.017)		(0.004)	(0.017)	
$\theta_s\tilde{\rho}_s \times \sum_{ij\in od} w_{ij}\mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$			0.094^a			0.108^a
			(0.015)			(0.015)
$\theta_s\times$ Pop-weighted Distance Groups	Y	Y	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y	Y	Y
Observations	712794	712794	712794	324289	324289	324289
# of Destination-Origin groups	2863	2863	2863	2790	2790	2790
1st-Stage F-stat		38.53			49.58	
R-squared	0.688		0.757	0.686		0.805

Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10. θ_s is from Caliendo and Parro (2015). Pop-weighted distance groups refer to the six population-weighted distance dummy variables.

Air shipments are skewed toward light, high-value products, so a value-based air share is systematically higher than a weight-based share, and excluding products by the value share drops more products at any cutoff; our baseline exclusion therefore uses the value share as the more conservative choice. As a check, we also compute the air share by weight and re-run the regressions.[33] Table B-5 shows that the results are similar to those in Table B-4, even though, at either cutoff, the weight-based samples retain considerably more observations.

[32] See https://usatrade.census.gov.

[33] The US Census data report import values shipped by air, by vessel, and in total, but weights only for air and vessel shipments. The weight share is therefore the air weight divided by the air plus vessel weight, while the value share divides by the total value, so the denominators of the two shares differ slightly. The smaller denominator inflates the weight share, so the comparison between the two shares is, if anything, conservative.

Table B-5: Robustness Checks: Exclude Products Transported by Air, Share $= \frac{\text{Air Weight}}{\text{Air Weight+Vessel Weight}}$

	Drop Products with Air Freight Share of above 5%			Drop Products with Air Freight Share of above 1%		
	OLS (1)	IV (2)	First Stage (3)	OLS (4)	IV (5)	First Stage (6)
Connection Index$_{od} \times \theta_s\tilde{\rho}_s$	0.025^a	0.034^b		0.024^a	0.050^a	
	(0.003)	(0.015)		(0.003)	(0.017)	
$\theta_s\tilde{\rho}_s \times \sum_{ij\in od} w_{ij}\mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$			0.091^a			0.093^a
			(0.015)			(0.015)
$\theta_s\times$ Pop-weighted Distance Groups	Y	Y	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y	Y	Y
Observations	1059443	1059443	1059443	512906	512906	512906
# of Destination-Origin groups	2897	2897	2897	2838	2838	2838
1st-Stage F-stat		36.66			37.49	
R-squared	0.689		0.741	0.686		0.756

Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10. θ_s is from Caliendo and Parro (2015). Pop-weighted distance groups refer to the six population-weighted distance dummy variables.

B.4.2 Alternative values of θ_s

We use θ_s estimated by Caliendo and Parro (2015) in our main specification. However, their estimates vary a lot across industries, with elasticities as high as 51.08 (Manufacturing of Petroleum Products) and sometimes large standard errors. To examine the robustness of our results with respect to the choice of θ_s, we first refer to the estimates in Giri et al. (2020). As discussed in footnote 9, Giri et al. (2020) use different data and variation to estimate θ_s, but both Giri et al. (2020) and Caliendo and Parro (2015) are consistent with the Eaton-Kortum model. The estimates in Giri et al. (2020) have a smaller variance across industries and tighter standard errors; however, they also cover fewer industries and result in a smaller sample.

Table B-6 shows the results using θ_s in Giri et al. (2020). Compared to our main results in Table 4, the sample sizes shrink by around 40%. However, the change in sample and θ_s does not affect our results: air connectedness positively affects trade in both the OLS and IV regressions. The estimates are higher than those using θ_s from Caliendo and Parro (2015).

Table B-6: Robustness Checks: Using θ_s from Giri et al. (2020)

	Full Sample	Sample in (5.5K,6.5K)		
	OLS	OLS	IV	First Stage
	(1)	(2)	(3)	(4)
Connection Index$_{od} \times \theta_s\tilde{\rho}_s$	0.028^a	0.041^a	0.045^b	
	(0.003)	(0.005)	(0.021)	
$\theta_s\tilde{\rho}_s \times \sum_{ji\in od} w_{ji}\mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$				0.092^a
				(0.015)
$\theta_s\times$ Pop-weighted Distance Groups	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y
Observations	3359848	1023749	1023749	1023749
# of Destination-Origin groups	9372	2878	2878	2878
1st-Stage F-stat			39.52	
R-squared	0.582	0.706		0.694

Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10. θ_s is from Giri et al. (2020). Pop-weighted distance groups refer to the six population-weighted distance dummy variables.

Given the uncertainty about the actual values of θ_s, we now consider a different strategy. We assume that θ_s is the same across sectors, an assumption maintained in Eaton and Kortum (2002). In this case, we can drop θ from the regressors and the main variable of interest becomes the product of air connectedness and $\tilde{\rho}_s$. Without a value of θ, we cannot interpret the coefficient of $\kappa_{od} \times \tilde{\rho}_s$ as the impact of air connectedness on communication costs m_{od}. However, we can still test whether the impact is significantly different from zero. Again, we see that $\kappa_{od} \times \tilde{\rho}_s$ has a positive effect on trade flows in both the OLS and IV regressions. The coefficient in the IV regression (Column (3)) is about 40% larger than that in the OLS regression. We can also take a stand on the value of θ. Suppose we use $\theta = 4$ from Simonovska and Waugh (2014), the IV estimate implies $\beta = 0.242/4 = 0.0605$, slightly larger than the IV estimate using θ_s from Giri et al. (2020).

Table B-7: Robustness Checks: Regressions without θ_s

	Full Sample	Sample in (5.5K,6.5K)		
	OLS	OLS	IV	First Stage
	(1)	(2)	(3)	(4)
Connection Index$_{od} \times \tilde{\rho}_s$	0.117^a	0.173^a	0.242^a	
	(0.015)	(0.019)	(0.089)	
$\tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$				0.095^a
				(0.015)
Pop-weighted Distance Groups	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y
Observations	5921180	1797671	1797671	1797671
# of Destination-Origin groups	9533	2928	2928	2928
1st-Stage F-stat			39.93	
R-squared	0.582	0.703		0.718

Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10. Pop-weighted distance groups refer to the six population-weighted distance dummy variables.

B.4.3 Alternative Instrumental Variables

In the main analysis, we select the share of grid pairs with distances between 5,500 and 6,500 miles that were distributed below 6,000 miles as an instrumental variable. In this section, we shrink the interval so that grid pairs falling below 6,000 miles within the interval are more likely to be "randomly assigned".

Table B-8 reports the results when using alternative windows when constructing z_{od}. The first Column reports the same IV estimation as in Column (3) of Table 4 for ease of comparison, and the next two columns show our robustness checks. When narrowing the window to (5.6K, 6.4K) and (5.7K,6.3K), we still find the IV estimates of β positive and statistically significant. However, the F-statistic becomes smaller when we use a narrower window to construct the IV. It is straightforward that when the window becomes narrower, it becomes harder to predict the overall air connectedness between two countries, especially for large countries that may have many grid pairs outside the window. We find that the IV would not significantly predict overall connectedness if we further narrowed the window.

Table B-8: Robustness Checks: Narrow the Window when Constructing the IV

IV Estimates	Sample in (5.5K,6.5K)		
	(5.5K,6.5K)	(5.6K,6.4K)	(5.7K,6.3K)
	(1)	(2)	(3)
Connection Index$_{od} \times \theta_s\tilde{\rho}_s$	0.026^b	0.047^b	0.053^b
	(0.012)	(0.019)	(0.026)
$\theta_s\times$ Pop-weighted Distance Groups	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y
Origin-HS6 FE	Y	Y	Y
First Stage	(1)	(2)	(3)
$\theta_s\tilde{\rho}_s \times \sum_{ij\in od} w_{ij}\mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$	0.091^a		
	(0.015)		
$\theta_s\tilde{\rho}_s \times \sum_{ij\in od} w_{ij}\mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.6K, 6.4K]\right)$		0.075^a	
		(0.019)	
$\theta_s\tilde{\rho}_s \times \sum_{ij\in od} w_{ij}\mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.7K, 6.3K]\right)$			0.072^a
			(0.026)
$\theta_s\times$ Pop-weighted Distance Groups	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y
Origin-HS6 FE	Y	Y	Y
Observations	1779937	1779937	1779937
# of Destination-Origin groups	2926	2926	2926
1st-Stage F-stat	36.29	15.51	7.77
R-squared	0.721	0.717	0.716

Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10. θ_s is from Caliendo and Parro (2015). Pop-weighted distance groups refer to the six population-weighted distance dummy variables.

In addition to adjusting the window when constructing the IVs, we also attempt to improve the first stage by using more than one instrumental variable. Campante and Yanagizawa-Drott (2018) show that connectedness between airports will have a network spillover effect: airports become more valuable with long-haul flights and also increase shorter-distance flights. Therefore, if two countries have more pairs of grids between 2,000 and 5,500 miles, our first instrumental variable, z_{od}, will have a larger impact on their connectedness. In Column (2) of Table B-9, we implement a specification using both $z_{od} \times \theta_s\tilde{\rho}_s$ and the product of $z_{od} \times \theta_s\tilde{\rho}_s$ and the share of grid pairs between 2,000 and 5,500 miles as IVs. As can be seen from the second panel (first stages), both variables positively predict air connectedness between countries. Column (3) uses the population-weighted share of grid pairs between 2,000 and 5,500 miles and finds similar results. However, the estimated β is very similar to that obtained using a single IV (Column (1)).

Table B-9: Robustness Checks: Multiple IVs with the First Stage Heterogeneity

IV Estimates	Single IV	Two IVs	
	(1)	(2)	(3)
Connection Index$_{od} \times \theta_s \tilde{\rho}_s$	0.026^b	0.022^b	0.022^b
	(0.012)	(0.010)	(0.010)
$\theta_s \times$ Pop-weighted Distance Groups	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y
Origin-HS6 FE	Y	Y	Y
First Stage	(1)	(2)	(3)
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$	0.091^a	0.064^a	0.072^a
	(0.015)	(0.024)	(0.022)
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$ $\times$ % in (2K,5.5K)		0.053^b	
		(0.025)	
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$ $\times$ Weighted % in (2K,5.5K)			0.040^c
			(0.021)
$\theta_s \times$ Pop-weighted Distance Groups	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y
Origin-HS6 FE	Y	Y	Y
Observations	1779937	1779937	1779937
# of Destination-Origin groups	2926	2926	2926
1st-Stage F-stat	36.29	33.29	33.30
R-squared	0.721	0.722	0.722

Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10.

B.4.4 Alternative Air Connectedness κ_{od}

In this section, we demonstrate that our results are robust to alternative methods for constructing the connection index κ_{od}. In our baseline, we treat two grids as connected if both fall within a 200-mile radius of airports that have weekly direct flights between them. We now vary the radius and assume the spillover effects of the connected airports can affect grids that are closer or farther away.

In Table B-10, we vary the radius r from 100 miles to 500 miles, including the baseline 200 miles reported in Column (3). In all columns, the IV estimates are positive and statistically significant. However, it is also clear that the estimates decline as the radius increases. One reason may be that the connected airports do not affect grids that are too far away. When we view those grids as "treated" and include them when calculating overall connectedness, we inflate the connectedness measure and obtain smaller estimates. We also report the average connection index under each radius selection. When an airport only affects cities within 100 miles, the mean connection index is 0.031. Intuitively, the average connection index between country pairs increases with radius selection, as countries are more likely to be connected if

airports can affect a wider range of cities.

Table B-10: Robustness Checks: Reduce or Increase Radius when Defining Connectedness

IV Estimates	$\log(value_{ods})$							
	$r = 100$ *miles*	$r = 150$ *miles*	$r = 200$ *miles*	$r = 250$ *miles*	$r = 300$ *miles*	$r = 350$ *miles*	$r = 400$ *miles*	$r = 500$ *miles*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Connection Index$_{od} \times \theta_s\tilde{\rho}_s$	0.068^b	0.035^b	0.026^b	0.019^b	0.016^b	0.014^b	0.013^b	0.012^b
	(0.031)	(0.016)	(0.012)	(0.009)	(0.007)	(0.007)	(0.006)	(0.006)
$\theta_s\times$ Pop-weighted Distance Groups	Y	Y	Y	Y	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y	Y	Y	Y	Y
First Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\theta_s\tilde{\rho}_s \times \sum_{ji\in od} w_{ji}\mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$	0.035^a	0.068^a	0.091^a	0.123^a	0.147^a	0.166^a	0.180^a	0.196^a
	(0.010)	(0.013)	(0.015)	(0.017)	(0.019)	(0.021)	(0.022)	(0.024)
$\theta_s\times$ Pop-weighted Distance Groups	Y	Y	Y	Y	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1779937	1779937	1779937	1779937	1779937	1779937	1779937	1779937
# of Destination-Origin groups	2926	2926	2926	2926	2926	2926	2926	2926
Mean Connection Index	0.031	0.051	0.067	0.086	0.103	0.117	0.128	0.145
1st-Stage F-stat	13.83	27.57	36.29	52.16	58.45	60.65	63.92	66.37
R-squared	0.690	0.714	0.721	0.735	0.743	0.746	0.749	0.756

Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10. θ_s is from Caliendo and Parro (2015).

In Table B-11, we experiment with the total number of direct flights required to determine whether two airports are connected. Instead of requiring a minimum of 52 round-trip direct flights, we increase the threshold to 104 (twice-weekly) and 365 (daily). The OLS and IV estimates are both significant and positive, and they increase in magnitude as the threshold is raised. This suggests that more frequent direct flights reduce communication costs further and lead to more trade flows.

Table B-11: Robustness Checks: Impact of More Frequent Direct Flights

	Twice-weekly Connection			Daily Connection		
	OLS	IV	First Stage	OLS	IV	First Stage
	(1)	(2)	(3)	(4)	(5)	(6)
Connection Index$_{od} \times \theta_s\tilde{\rho}_s$	0.024^a	0.031^b		0.028^a	0.048^b	
	(0.003)	(0.014)		(0.003)	(0.022)	
$\theta_s\tilde{\rho}_s \times \sum_{ji\in od} w_{ji}\mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$			0.078^a			0.050^a
			(0.015)			(0.015)
$\theta_s\times$ Pop-weighted Distance Groups	Y	Y	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y	Y	Y
Observations	1779937	1779937	1779937	1779937	1779937	1779937
# of Destination-Origin groups	2926	2926	2926	2926	2926	2926
1st-Stage F-stat		25.38			11.60	
R-squared	0.700		0.710	0.700		0.629

Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10. θ_s is from Caliendo and Parro (2015). Pop-weighted distance groups refer to the six population-weighted distance dummy variables. A twice-weekly (or daily) connection is defined as having more than 104 (or 365) round-trip direct flights between two countries in a year.

B.4.5 Remove the Country Boundaries within the Schengen Zone

Schengen Zone is a group of 26 (as of 2016) European countries that have abolished passport and other types of border control at their mutual borders. Therefore, it is reasonable to assume that passengers can easily reach a nearby city in the neighboring country to catch a flight. When determining κ_{ji}, we removed national borders within the Schengen Zone, which means the 200-mile spillover effect applies across any Schengen country, and subsequently aggregated the value at the country-pair level.[34] Table B-12 and Figure B-3 present that removing internal borders within the Schengen Zone increased the average connection index by 30%. The Schengen index is highly correlated with the original index, with a correlation coefficient of 0.99.

Table B-13 reports the results when we remove the country boundaries within the Schengen Zone when calculating the connection index κ_{od}. The results are similar to our main results in Table 4.

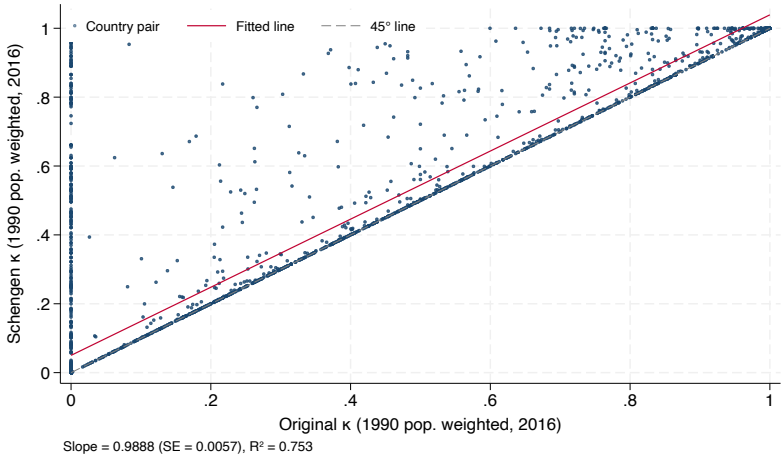
Table B-12: Summary Statistics: Original vs. Schengen Connection Index (2016)

	Obs.	mean	Std. Dev.	p25	p50	p75	p99	max
Original Index	9900	0.158	0.298	0.000	0.000	0.174	1.000	1.000
Schengen Index	9900	0.207	0.340	0.000	0.000	0.334	1.000	1.000
Difference (Schengen − Original)	9900	0.048	0.169	0.000	0.000	0.000	0.989	1.000

Notes. Summary statistics for the population-weighted country-pair connection index in 2016. The original index respects national borders when determining which grids are served by an airport. The Schengen index removes internal Schengen borders, allowing any grid within 200 miles of an airport to be counted as served regardless of which member state the airport is in. The difference is non-negative by construction.

[34] Note that we do not treat the Schengen Zone as a single country: we still use the population within each country's grid cells as weights to calculate the connectedness between that country and others. However, at the grid level—once national borders within the Schengen Zone are removed—the probability of direct flights existing between two grids increases.

Figure B-3: Original vs. Schengen Connection Index, Country Pairs (2016)



Notes. Each point is a country pair. The horizontal axis is the original connection index and the vertical axis is the Schengen index. The dashed line is the 45° line; the solid red line is the OLS fit.

Table B-13: Robustness Checks: Impact of Removing the Country Boundaries within the Schengen Zone

	Full Sample	Sample in (5.5K,6.5K)		
	OLS	OLS	IV	First Stage
	(1)	(2)	(3)	(4)
Connection Index$_{od} \times \theta_s\tilde{\rho}_s$	0.014^a	0.019^a	0.025^b	
	(0.002)	(0.002)	(0.011)	
$\theta_s\tilde{\rho}_s \times \sum_{ji\in od} w_{ji}\mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$				0.097^a
				(0.015)
$\theta_s\times$ Pop-weighted Distance Groups	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y
Observations	5855591	1779937	1779937	1779937
# of Destination-Origin groups	9524	2926	2926	2926
1st-Stage F-stat			43.55	
R-squared	0.563	0.699		0.791

Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10. θ_s is from Caliendo and Parro (2015). Pop-weighted distance groups refer to the six population-weighted distance dummy variables. κ_{od} is calculated by removing the country boundaries within the Schengen Zone when finding the neighboring airports.

B.4.6 Using Road Distance to Construct κ_{od}

When calculating our baseline connectedness index, we define the catchment area of an airport using the great circle distance between grids. In this section, we conduct a robustness check using road distance instead. We use the Fast Marching Method (FMM) to calculate a "normalized road distance" and then use this distance to determine whether an airport is

accessible for each grid. The road network comes from the OpenStreetMap project; we keep the major highway classes and the reachable waterways.[35] Following Allen and Arkolakis (2014), we convert the network into a cost raster at 2 kilometer resolution under an Azimuthal Equidistant projection centered on each country's centroid. Each pixel is assigned a traversal cost inversely proportional to assumed travel speed. We normalize the cost of "secondary" roads, the slowest highway class, to 1; the cost is then $\frac{20}{70}$ for "motorway" roads, $\frac{20}{55}$ for "trunk" roads, and $\frac{20}{35}$ for "primary" roads. Accessible waterways receive a cost of $\frac{20}{15}$; roads outside the highway classes and areas without road data receive $\frac{20}{3}$.[36] Table B-14 compares the two distance measures across within-country grid pairs. Road distance is in general larger than great circle distance, and the gap can be large. However, the summary statistics cover grid pairs at all distances. For an airport's catchment area, only short distances matter, and at short distances the two measures are close.

Table B-14: Summary Statistics: Distance in Miles

	Obs.	mean	std. dev.	min	p1	p5	p10	p25	p50	p75	p90	p95	p99	max
Great Circle Distance (Miles)	23,089,516	1,477	911	0	119	284	430	778	1,322	2,003	2,760	3,253	4,041	10,951
Normalized Road Distance (Miles)	23,089,496	2,842	2,332	0	168	373	557	1,118	2,347	3,951	5,745	7,075	9,527	97,545

Notes. The number of observations for road distance is 20 fewer than the number for great circle distance because the road network data for Trinidad and Tobago is missing. Trinidad and Tobago have 4 grids, which can form a total of 4 × 5=20 pairs.

Table B-15 shows the robustness check for how connectedness affects trade after calculating the κ_{od} using road distance obtained through the FMM algorithm. Comparing with Table B-10, we can see that when choosing similar thresholds, the regression results obtained using road distance and great circle distance are highly consistent: at the same threshold, higher connectedness leads to more trade, and this effect gradually decreases as we increase the radius that the airport can reach. The regression results using road distance are slightly lower than those using great circle distance at the same threshold.

[35]The road network data is provided by https://www.geofabrik.de. The major highway class includes motorway, trunk, primary, and secondary, which are ordered from the highest speed to the lowest speed (from the lowest cost to the highest cost). The reachable waterway represents ferry service for people and transportation between different ports. Ferry service has a cost slightly higher than highways. Their data update daily and we use the version of 7 August 2026.

[36]Our rasterization follows the logic of "lowest cost wins." We rasterize layer by layer according to cost, from highest to lowest. In each 2km × 2km pixel, if a certain road is found, we assign a corresponding cost to that pixel. However, if a lower-cost road is found in a later layer, the lower-cost road replaces the highest-cost road, ultimately ensuring that each pixel has a potential lowest cost. This might lead to some pixels with mixed roads receiving lower costs, resulting in overestimating the actual travel speed of that road. However, considering that a 2km pixel is relatively small for a road, and that we also impose a very high cost penalty on roads other than major highways, the overestimation problem will not be too serious. Ferry speeds are sourced from the Bureau of Transportation Statistics: https://www.bts.gov/surveys/national-census-ferry-operators-ncfo/ncfo-previous-datasets. These statistics show that the average ferry speed is 13.5 knots, which is approximately 15 mph. For speeds in blank areas, average walking speed is used.

Table B-15: Robustness Checks: Using Road Distance to Construct κ_{od}

IV Estimates	$\log(value_{ods})$							
	$r = 100$ *miles*	$r = 150$ *miles*	$r = 200$ *miles*	$r = 250$ *miles*	$r = 300$ *miles*	$r = 350$ *miles*	$r = 400$ *miles*	$r = 500$ *miles*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Connection Index$_{od} \times \theta_s \tilde{\rho}_s$	0.058^b	0.032^b	0.022^b	0.017^b	0.016^b	0.015^b	0.014^b	0.012^b
	(0.029)	(0.015)	(0.011)	(0.008)	(0.007)	(0.007)	(0.006)	(0.005)
$\theta_s \times$ Pop-weighted Distance Groups	Y	Y	Y	Y	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y	Y	Y	Y	Y
First Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\theta_s \tilde{\rho}_s \times \sum_{ji \in od} w_{ji} \mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right)$	0.041^a	0.074^a	0.108^a	0.146^a	0.152^a	0.158^a	0.173^a	0.199^a
	(0.007)	(0.011)	(0.014)	(0.017)	(0.021)	(0.022)	(0.024)	(0.026)
$\theta_s \times$ Pop-weighted Distance Groups	Y	Y	Y	Y	Y	Y	Y	Y
Dest.-HS6 FE	Y	Y	Y	Y	Y	Y	Y	Y
Origin-HS6 FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1779937	1779937	1779937	1779937	1779937	1779937	1779937	1779937
# of Destination-Origin groups	2926	2926	2926	2926	2926	2926	2926	2926
Mean Connection Index	0.023	0.047	0.069	0.094	0.119	0.139	0.151	0.170
1st-Stage F-stat	36.22	43.28	56.51	73.06	54.34	50.47	53.17	59.13
R-squared	0.719	0.739	0.754	0.762	0.762	0.770	0.766	0.764

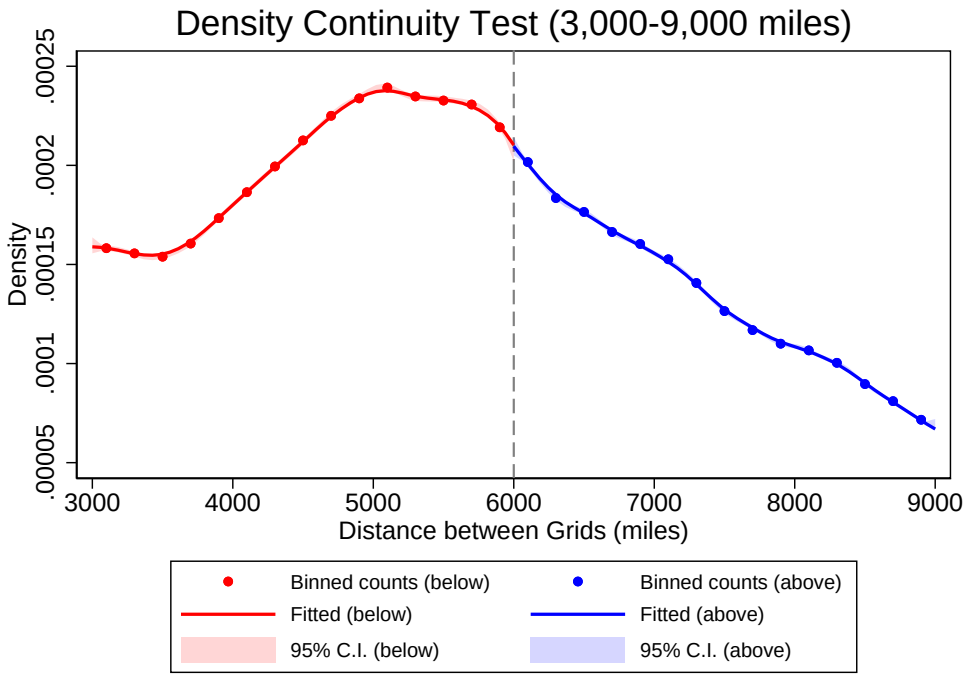
Notes. Standard errors are clustered at the destination-origin country pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10. θ_s is from Caliendo and Parro (2015).

B.5 RD Diagnostic Tests

This section reports standard diagnostics for the grid-pair regression discontinuity design in Section 5.2: a density test on the running variable and balance tests on predetermined covariates.

If it were known beforehand that routes within 6,000 miles are significantly cheaper to operate than those exceeding 6,000 miles, the distribution of airport locations might be artificially biased towards areas within 6,000 miles, making grid pairs just above and just below the cutoff incomparable. To rule this out, we perform a density test on the running variable. Figure B-4 shows that the number of grid pairs is continuous at the 6,000-mile threshold, with no significant discontinuity in the density at the cutoff. This evidence suggests that the distribution of airport locations is not artificially biased towards areas within 6,000 miles.

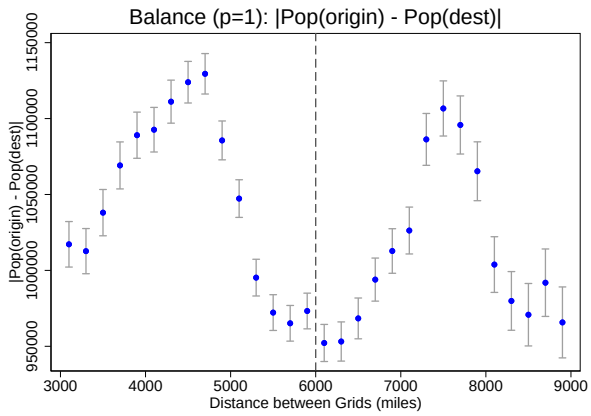
Figure B-4: RD Diagnostic: Density Test for the Running Variable



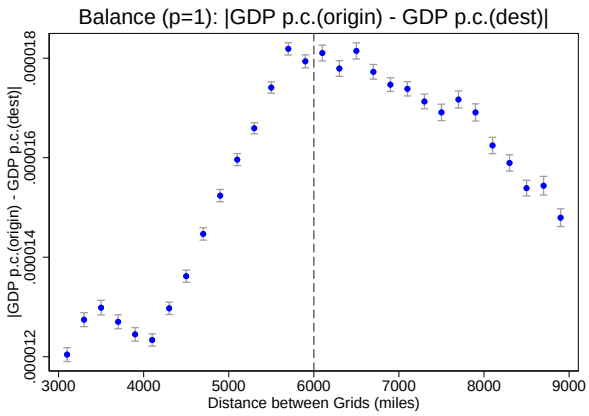
Notes. This figure presents the standard McCrary (2008) density continuity test for the running variable. Each dot displays the total counts of grid-pair observations in each 200-mile bin from 3,000 to 9,000 miles, with the cutoff at 6,000 miles. The solid curves are local polynomial density estimates fitted separately on each side of the cutoff, and the shaded areas represent 95% confidence intervals (default choices in Stata's `rddensity` command).

We also perform balance tests on predetermined economic and geographic variables. We calculate the absolute differences between the two grids in population, GDP per capita, and coordinates, and plot how their means change with distance. All four panels of Figure B-5 show that these covariates have no sharp change at the 6,000-mile threshold, consistent with the RD assumption that covariates are continuous at the cutoff of the running variable.

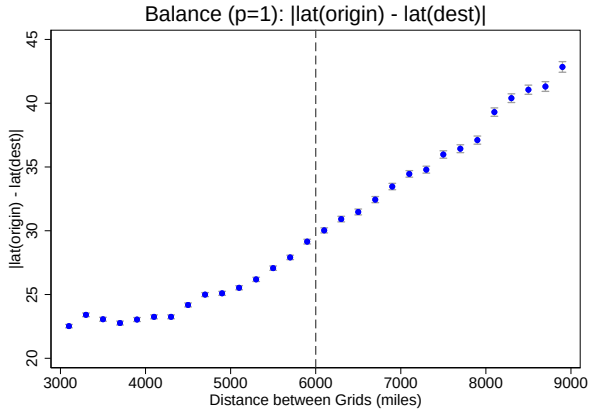
Figure B-5: RD Diagnostic: Balance Test for Covariates at the 6,000-Mile Threshold



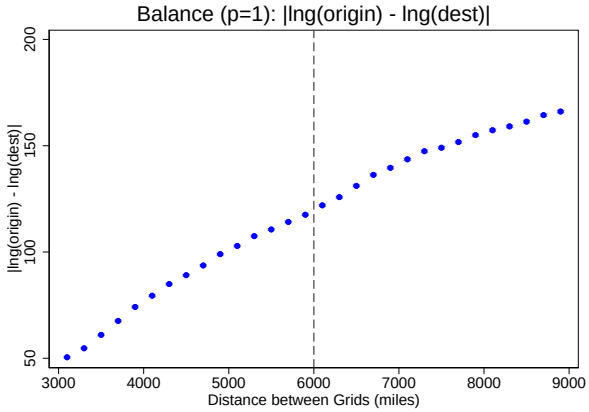
(a) |Pop(origin) − Pop(dest)|



(b) |GDP p.c.(origin) − GDP p.c.(dest)|



(c) |lat(origin) − lat(dest)|



(d) |lng(origin) − lng(dest)|

Notes. Standard RD diagnostic for covariates. Each panel plots the absolute difference of a grid-level characteristic between origin and destination grids against the distance between them. The running variable is the distance between grid pairs (miles) and the cutoff is 6,000 miles. There are 15 bins on each side of the cutoff. The dots represent the average covariate difference within each bin. The bar indicates the 95% confidence intervals, calculated based on the standard errors from a global first-order polynomial regression using data from 3,000 to 9,000 miles; they are for visualization purposes only and differ from the RD analysis conducted in the main text. Grid-level population and GDP data are from the G-Econ dataset. GDP per capita is calculated as GDP divided by population.

B.6 Robustness of the impact of distance on business lawsuits and foreign ownership links

Table B-16: Sharp RD: Effects of Distance on Number of Cases and Foreign Ownership Links at the Grid-pair Level, Uniform Kernel

Dep. Var.:	# DP Cases$_{ji}$ (×10)				# FDI Links$_{ji}$			
Bandwidth (Miles):	500	800	Optimal		500	800	Optimal	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}(d_{ji} \leq 6K$ miles)	-0.0000	0.0000	-0.0002	-0.0000	0.0317^{c}	0.0383^{a}	0.0326^{b}	0.0229
	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0170)	(0.0119)	(0.0165)	(0.0213)
Observations	2582388	2582388	2582388	2582388	2582388	2582388	2582388	2582388
Observations, effective	355730	561598	439848	717184	355730	561598	368406	573652
Polynomial Order	1	1	1	2	1	1	1	2
Kernel Choice	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform
Bandwidth	500	800	621	1029	500	800	519	818
Mean of Dep. Var.	0.0006	0.0007	0.0007	0.0007	0.0421	0.0417	0.0418	0.0413

Notes. Same as Table 6 but using a uniform kernel instead of a triangular kernel. Significance levels: a: 0.01, b: 0.05, c: 0.10.

Table B-17: Sharp RD: Effects of Distance on Number of Cases and Foreign Ownership Links at the Grid-pair Level: Excluding USA from the Sample

Dep. Var.:	10 × # Defective Product Case$_{ji}$				# Foreign Ownership Links$_{ji}$			
Bandwidth (Miles):	500	800	Optimal		500	800	Optimal	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}(d_{ji} \leq 6K$ miles)	0.0001	0.0002	0.0002	0.0002	0.0299	0.0362^{c}	0.0364^{c}	0.0365
	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0297)	(0.0211)	(0.0198)	(0.0263)
Observations	1955632	1955632	1955632	1955632	1955632	1955632	1955632	1955632
Observations, effective	252588	394470	635704	759508	252588	394470	425140	575290
Polynomial Order	1	1	1	2	1	1	1	2
Kernel Choice	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Bandwidth	500	800	1350	1670	500	800	866	1205
Mean of Dep. Var.	0.0002	0.0001	0.0001	0.0001	0.0392	0.0325	0.0311	0.0273

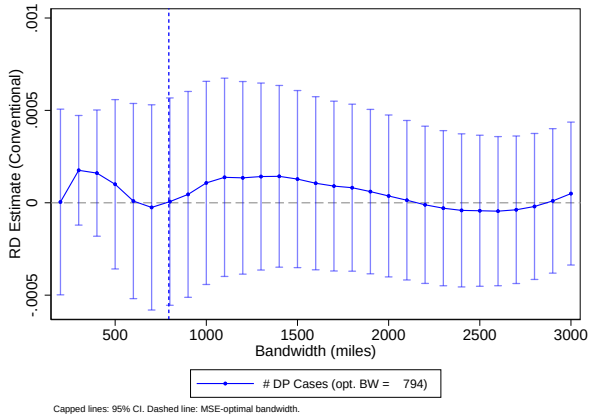
Notes. Local polynomial regression discontinuity estimates. We drop all cases or foreign ownership links involving the USA. The dependent variables are the number of business lawsuits and foreign ownership links. When the dependent variable is the number of litigation cases, the values are scaled up 10 times due to the small number of business lawsuit cases. The running variable is the distance between grids, discontinuity at 6,000 miles. The optimal bandwidth is selected using the MSE-optimal bandwidth selector, and the triangular kernel function is adopted (both are default choices in Stata's `rdrobust` command). "Observations" refer to the total number of grid pairs across all distance ranges, while "effective observations" only include grid pairs within the bandwidth. Robust standard errors are in parentheses. Significance levels: a: 0.01, b: 0.05, c: 0.10.

Table B-18: Sharp RD: Effects of Distance on Number of Cases at the Grid-pair Level, Control for Population, GDP per Capita and Coordinate Differences

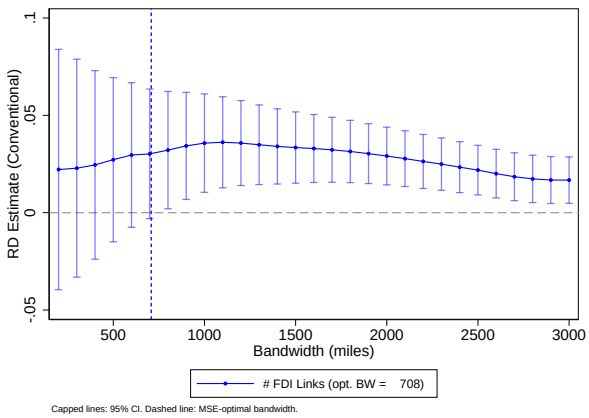
Dep. Var.:	# DP Cases$_{ji}$ (×10)				# FDI Links$_{ji}$			
Bandwidth (Miles):	500	800	Optimal		500	800	Optimal	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}(d_{ji} \leq 6K$ miles)	0.0002	0.0000	0.0000	0.0000	0.0354	0.0403^{c}	0.0386^{c}	0.0418^{c}
	(0.0003)	(0.0004)	(0.0004)	(0.0003)	(0.0291)	(0.0208)	(0.0220)	(0.0249)
Observations	1882438	1882438	1882438	1882438	1882438	1882438	1882438	1882438
Observations, effective	260520	411352	415238	460406	260520	411352	381632	643696
Control	Econ, Geo	Econ, Geo	Econ, Geo	Econ, Geo	Econ, Geo	Econ, Geo	Econ, Geo	Econ, Geo
Polynomial Order	1	1	1	2	1	1	1	2
Kernel Choice	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Bandwidth	500	800	808	897	500	800	740	1274
Mean of Dep. Var.	0.0006	0.0007	0.0007	0.0008	0.0421	0.0417	0.0417	0.0378

Notes. Local polynomial regression discontinuity estimates. The dependent variables are the number of "defective product" business lawsuits and foreign ownership links. The absolute differences between the two grids in population, GDP per capita, latitude, and longitude are included as control variables. Significance levels: a: 0.01, b: 0.05, c: 0.10.

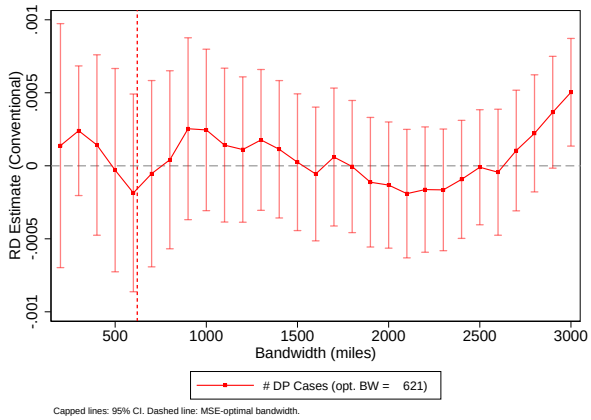
Figure B-6: Coefficients Comparison across a Range of Bandwidths



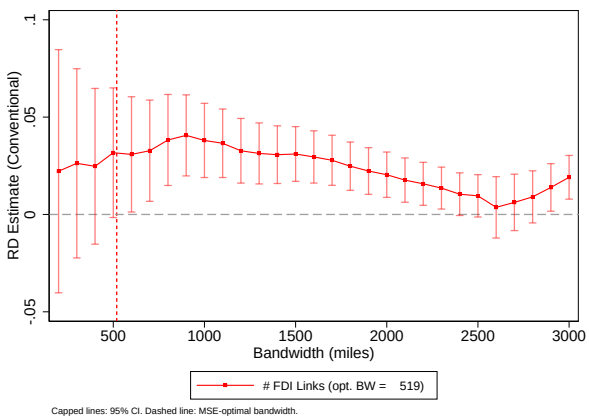
(a) # Cases, Triangular Kernel



(b) # FDI Links, Triangular Kernel



(c) # Cases, Uniform Kernel



(d) # FDI Links, Uniform Kernel

Notes. This figure shows the coefficients of first-order sharp RD regressions as the bandwidth varies from 200 miles to 3,000 miles. Panels (a) and (b) show the effects of distance on both outcomes when using a triangular kernel, while panels (c) and (d) show the effects when using a uniform kernel. The optimal bandwidth is marked with a dashed line.

Table B-19: Sharp RD: Effects of Distance on Number of Cases and Foreign Ownership Links, Outcome Variables are Transformed as $\text{arcsinh}(Y_{ji})$

Dep. Var.:	arcsinh(# Defective Product Case$_{ji}$)			arcsinh(Revenue$_{ji}$)			arcsinh(# Foreign Ownership Links$_{ji}$)		
Bandwidth (Miles):	500	800	Optimal	500	800	Optimal	500	800	Optimal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\mathbb{1}(d_{ji} \leq 6K$ miles)	0.0001	0.0000	0.0000	0.0328^a	0.0436^a	0.0283^a	0.0042^a	0.0060^a	0.0031^b
	(0.0001)	(0.0001)	(0.0001)	(0.0078)	(0.0061)	(0.0090)	(0.0013)	(0.0010)	(0.0015)
Observations	2582388	2582388	2582388	2582388	2582388	2582388	2582388	2582388	2582388
Observations, effective	355730	561598	479076	355730	561598	266770	355730	561598	263954
Polynomial Order	1	1	1	1	1	1	1	1	1
Kernel Choice	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Bandwidth	500	800	679	500	800	374	500	800	370
Mean of Dep. Var.	0.0004	0.0004	0.0004	0.0962	0.0947	0.0960	0.0132	0.0129	0.0132

Notes. The dependent variables are the number of business lawsuits, the total revenue of foreign subsidiaries, and foreign ownership links transformed by the $\text{arcsinh}(y) = \log(y + \sqrt{y^2 + 1})$ function. The running variable is the distance between grids, discontinuity at 6,000 miles. The optimal bandwidth is selected using the MSE-optimal bandwidth selector, and the triangular kernel function is adopted (both are default choices in Stata's `rdrobust` command). "Observations" refer to the total number of grid pairs across all distance ranges, while "effective observations" only include grid pairs within the bandwidth. Robust standard errors are in parentheses. Significance levels: a: 0.01, b: 0.05, c: 0.10.

Table B-20: Sharp RD: Effects of Distance on Number of Cases and Foreign Ownership Links, Outcome Variables are Transformed as $\log(Y_{ji} + 1)$

Dep. Var.:	log(# Defective Product Case$_{ji}$ + 1)			log(Revenue$_{ji}$+1)			log(# Foreign Ownership Links$_{ji}$+1)		
Bandwidth (Miles):	500	800	Optimal	500	800	Optimal	500	800	Optimal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\mathbb{1}(d_{ji} \leq 6K$ miles)	0.0000	0.0000	0.0000	0.0308^a	0.0411^a	0.0265^a	0.0033^a	0.0048^a	0.0025^b
	(0.0001)	(0.0001)	(0.0001)	(0.0073)	(0.0057)	(0.0085)	(0.0010)	(0.0008)	(0.0012)
Observations	2582388	2582388	2582388	2582388	2582388	2582388	2582388	2582388	2582388
Observations, effective	355730	561598	479436	355730	561598	266228	355730	561598	265174
Polynomial Order	1	1	1	1	1	1	1	1	1
Kernel Choice	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Bandwidth	500	800	680	500	800	373	500	800	371
Mean of Dep. Var.	0.0003	0.0003	0.0003	0.0902	0.0888	0.0898	0.0104	0.0102	0.0104

Notes. The dependent variables are the number of business lawsuits, the total revenue of foreign subsidiaries, and foreign ownership links transformed by the $\log(y + 1)$ function. The running variable is the distance between grids, discontinuity at 6,000 miles. The optimal bandwidth is selected using the MSE-optimal bandwidth selector, and the triangular kernel function is adopted (both are default choices in Stata's `rdrobust` command). "Observations" refer to the total number of grid pairs across all distance ranges, while "effective observations" only include grid pairs within the bandwidth. Robust standard errors are in parentheses. Significance levels: a: 0.01, b: 0.05, c: 0.10.

Table B-21: OLS: Effects of Distance Discontinuity Index on Number of Defective Product Cases, Controlling for FDI Links

Dep. Var.:	log(# DP Cases$_{od}$)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right.$ miles)	-1.199^{a}	-0.819	-0.927^{b}	-0.723	-1.295^{a}	-0.784
	(0.414)	(0.561)	(0.466)	(0.654)	(0.428)	(0.563)
log(# FDI Links$_{od}$)	0.593^{a}	0.759^{b}				
	(0.223)	(0.296)				
$\mathbb{1}$(# FDI $\leq$ median$\mid$# FDI $>$ 0)			-0.046	0.176		
			(0.910)	(0.959)		
$\mathbb{1}$(# FDI $>$ median$\mid$# FDI $>$ 0)			0.464	0.799		
			(0.913)	(0.956)		
log(1 + # FDI Links$_{od}$)					0.675^{a}	0.806^{a}
					(0.216)	(0.253)
Control: Pop-weighted Distance Groups	N	Y	N	Y	N	Y
Origin-Country FE	Y	Y	Y	Y	Y	Y
Dest.-Country FE	Y	Y	Y	Y	Y	Y
Observations	157	157	171	171	171	171
# of Destination-Origin groups	157	157	171	171	171	171
R-squared	0.856	0.867	0.833	0.840	0.850	0.859

Notes. OLS regressions at the country-pair level. The dependent variables are the log-transformed number of defective product lawsuits at the country-pair level. The key regressor is the pop-weighted share of grid pairs within a country pair whose distance falls just below 6,000 miles (within the 5,500-6,500 mile window). FDI links are controlled for either continuously (log transformation) or via decile group indicators. Origin- and destination-country fixed effects are included in all columns. Standard errors are clustered at the country-pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10.

Table B-22: PPML: Effects of Distance Discontinuity Index on Number of Defective Product Cases, Controlling for FDI Links

Dep. Var.:	# DP Cases$_{od}$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K | d_{ij} \in (5.5K, 6.5K]\right.$ miles)	0.018	-0.408	0.419	-0.344	-0.079	-0.538
	(0.314)	(0.510)	(0.391)	(0.546)	(0.289)	(0.465)
log(# FDI Links$_{od}$)	0.612^{a}	0.696^{a}				
	(0.133)	(0.228)				
$\mathbb{1}$ (# FDI $\leq$ median$\vert$# FDI $>$ 0)			1.203^{b}	1.200^{c}		
			(0.596)	(0.615)		
$\mathbb{1}$ (# FDI $>$ median$\vert$# FDI $>$ 0)			2.444^{a}	2.280^{a}		
			(0.624)	(0.633)		
log(1 + # FDI Links$_{od}$)					0.566^{a}	0.549^{a}
					(0.117)	(0.172)
Control: Pop-weighted Distance Groups	N	Y	N	Y	N	Y
Origin-Country FE	Y	Y	Y	Y	Y	Y
Dest.-Country FE	Y	Y	Y	Y	Y	Y
Observations	566	566	864	864	864	864
# of Destination-Origin groups	566	566	864	864	864	864
pseudo-R-squared	0.891	0.892	0.900	0.901	0.900	0.901

Notes. PPML regressions at the country-pair level. The dependent variable is the number of defective product lawsuits at the country-pair level. The key regressor is the pop-weighted share of grid pairs within a country pair whose distance falls just below 6,000 miles (within the 5,500-6,500 mile window). FDI links are controlled for either continuously (log transformation) or via decile group indicators. Origin- and destination-country fixed effects are included in all columns. Standard errors are clustered at the country-pair level. Significance levels: a: 0.01, b: 0.05, c: 0.10.

B.7 Results Using "Breach of Contract" Cases

As discussed in Section 3 (footnote 7), we also extract business lawsuits using the keyword "breach of contract", which captures business disputes in general. Many of these cases involve services or disputes unrelated to manufactured products, so we do not use them in the main analysis; this section reports the corresponding results as a robustness check. We process the cases in the same way as the defective-product cases, keeping firm-vs-firm cases and matching the parties to global ultimate owners in ORBIS. Excluding cases between firms from the same country, we obtain 9,480 breach-of-contract lawsuits between firms, distributed across 1,937 country pairs; 269 cases contain both keywords. Table B-23 lists the 30 country pairs with the most cases.

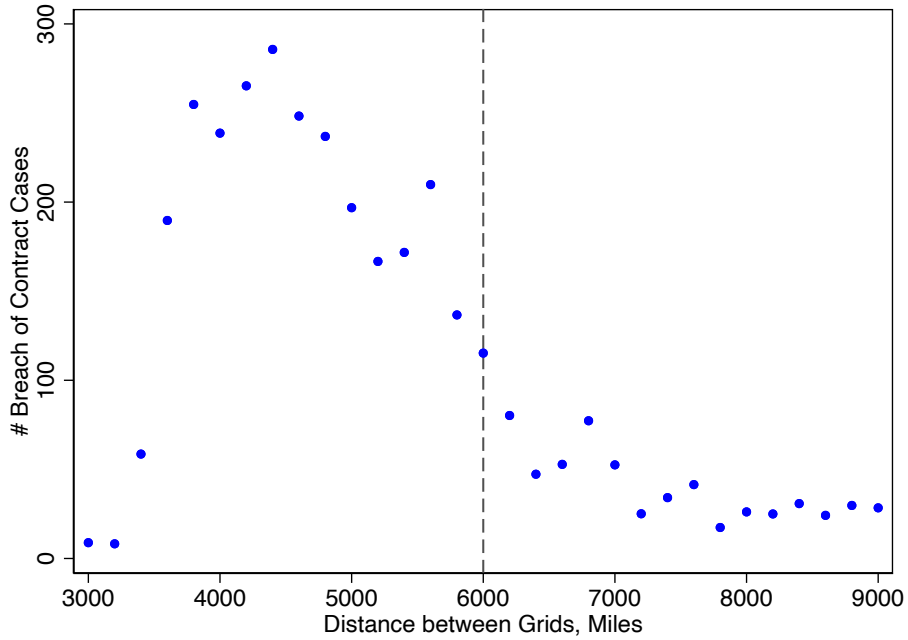
Table B-23: List of Country-pairs with the Highest Number of Cases, Breach of Contract

Rank	Country Pair	# Litigation Cases	Rank	Country Pair	# Litigation Cases
1	USA-DEU	404	16	USA-FRA	121
2	DEU-USA	391	17	BEL-USA	72
3	USA-ITA	386	18	USA-BEL	59
4	ITA-USA	376	19	RUS-USA	55
5	GBR-USA	354	20	AUT-USA	47
6	USA-AUS	343	21	USA-RUS	46
7	USA-GBR	334	22	USA-AUT	39
8	AUS-USA	308	23	SWE-USA	38
9	USA-NLD	286	24	CHN-USA	34
10	NLD-USA	279	25	USA-ISR	34
11	CAN-USA	240	26	USA-ESP	34
12	USA-CAN	195	27	ESP-USA	33
13	JPN-USA	142	28	IND-USA	32
14	FRA-USA	128	29	ISR-USA	31
15	USA-JPN	124	30	USA-SWE	31

Notes. The Lexis US data includes 1,937 country pairs with some "breach of contract" lawsuits. Here, we list the top 30 country pairs with the highest number of such lawsuits. The first country refers to the location of the plaintiffs, while the second country refers to the location of the defendants. The average number of cases between countries is 4 when searching for "breach of contract". In 1,770 country pairs, both the plaintiffs and the defendants are located outside the US.

Figure B-7 plots the number of breach-of-contract cases between grid pairs by distance. As with the defective-product cases, the number of cases declines smoothly through the 6,000-mile threshold, with no visible discontinuity at the cutoff.

Figure B-7: Number of "Breach of Contract" Cases between Grid Pairs, by Distance



Notes. We group grid pairs into 200-mile bins. Each dot represents the number of "breach of contract" cases in the preceding 200 miles. The sample of grid pairs is the same as in Figure 2.

Table B-24 repeats the sharp RD analysis of Section 5.2 with the breach-of-contract cases as the outcome. Consistent with the defective-product results, having a distance of less than 6,000 miles does not significantly affect the number of cases, while the number of foreign ownership links jumps significantly at the cutoff.

Table B-24: Sharp RD: Effects of Distance on Number of Cases and Foreign Ownership Links at the Grid-pair Level, "Breach of Contract" as Keyword

Dep. Var.:	# BoC Cases$_{ji}$ (×10)				# FDI Links$_{ji}$			
Bandwidth (Miles):	500	800	Optimal		500	800	Optimal	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}(d_{ji} \leq 6K$ miles)	0.0001	0.0013	-0.0000	-0.0022	0.0272	0.0321^b	0.0303^c	0.0327^c
	(0.0026)	(0.0022)	(0.0024)	(0.0030)	(0.0215)	(0.0154)	(0.0168)	(0.0186)
Observations	2582388	2582388	2582388	2582388	2582388	2582388	2582388	2582388
Observations, effective	355730	561598	449578	594170	355730	561598	499696	867664
Polynomial Order	1	1	1	2	1	1	1	2
Kernel Choice	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Bandwidth	500	800	635	847	500	800	708	1259
Mean of Dep. Var.	0.0141	0.0159	0.0150	0.0158	0.0421	0.0417	0.0420	0.0377

Notes. Local polynomial regression discontinuity estimates. The dependent variables are the number of "breach of contract" business lawsuits and foreign ownership links. When the dependent variable is the number of litigation cases, the values are scaled up 10 times due to the small number of business lawsuit cases. The running variable is the distance between grids, discontinuity at 6,000 miles. The optimal bandwidth is selected using the MSE-optimal bandwidth selector, and the triangular kernel function is adopted (both are default choices in Stata's `rdrobust` command). "Observations" refer to the total number of grid pairs across all distance ranges, while "effective observations" only include grid pairs within the bandwidth. Robust standard errors are in parentheses. Significance levels: a: 0.01, b: 0.05, c: 0.10.

C Details of the Hat Algebra

We now derive some key equations for the "exact hat algebra". We express the change in labor demand in terms of changes in wages across all countries. We then solve for the wage changes from the labor market clearing conditions and predict the welfare changes resulting from altering the communication costs, m_{od}, to m'_{od}. We have specified earlier that $m_{od} = e^{\beta_0 - \beta\kappa_{od}}$. Therefore,

$$\frac{m'_{od}}{m_{od}} = e^{-\beta\left(\kappa'_{od} - \kappa_{od}\right)}, \tag{C-1}$$

where κ_{od} and κ'_{od} are the connectedness index between countries o and d in the base and later periods, respectively.

Denoting $\hat{x} \equiv x'/x$ where x' is the outcome after the shocks, we can write the price equation (6) in changes:

$$\hat{P}_{ds}^{-\theta_s} = \sum_o \pi_{ods} \left(\hat{m}_{od}^{\tilde{\rho}_s} \hat{c}_{os}\right)^{-\theta_s}. \tag{C-2}$$

The change in the cost of the input bundle, $\hat{c}_{os}$, can be obtained from equation (12):

$$\hat{c}_{os} = \hat{w}_o^{a_{L,os}} \prod_k \hat{P}_{ok}^{a_{ok,s}}. \tag{C-3}$$

Combining $2N \times S$ nonlinear equations (C-2) and (C-3), we can solve $\hat{P}_{os}, \hat{c}_{os}$ as functions

of $\hat{w}_o$, $\hat{m}_{od}$ and base-period shares. We also obtain the change in trade shares as

$$\hat{\pi}_{ods} = \left[\frac{T_{os}\left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od}\right)^{-\theta_s}}{\sum_o T_{os}\left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od}\right)^{-\theta_s}}\right]^{-1} \frac{T_{os}\left((m'_{od})^{\tilde{\rho}_s} c'_{os} t_{od}\right)^{-\theta_s}}{\sum_o T_{os}\left((m'_{od})^{\tilde{\rho}_s} c'_{os} t_{od}\right)^{-\theta_s}} = \frac{\left(\hat{m}_{od}^{\tilde{\rho}_s} \hat{c}_{os}\right)^{-\theta_s}}{\sum_o \pi_{ods}\left(\hat{m}_{od}^{\tilde{\rho}_s} \hat{c}_{os}\right)^{-\theta_s}} \tag{C-4}$$

The goods market clearing conditions (13) can be written in changes as

$$\hat{R}_{os} = \sum_d \frac{\pi_{ods}\xi_{ds}E_d}{R_{os}} \hat{\pi}_{ods}\hat{E}_d + \sum_{d,k} \frac{\pi_{ods}a_{os,k}R_{dk}}{R_{os}} \hat{\pi}_{ods}\hat{R}_{dk}, \tag{C-5}$$

where the change in expenditure $\hat{E}_d$ is

$$\hat{E}_d = \frac{w_d L_d}{E_d}\hat{w}_d + \frac{D'_d}{E_d}.$$

D'_d denotes the exogenous trade deficits in the counterfactual equilibrium. In our application they scale with income, $D'_d = \hat{w}_d D_d$, so that $\hat{E}_d = \hat{w}_d$ by the identity $E_d = w_d L_d + D_d$ (see footnote 17 and Online Appendix D.3 for the alternative convention that holds D_d fixed). Similar to equation (13), equation (C-5) is a linear system of $\hat{R}_{os}$, and we can apply matrix inversion to obtain the solution.

The labor market clearing conditions (15) can be written in changes as

$$\hat{w}_o = \sum_s \frac{a_{L,os}R_{os}}{w_o L_o}\hat{R}_{os}. \tag{C-6}$$

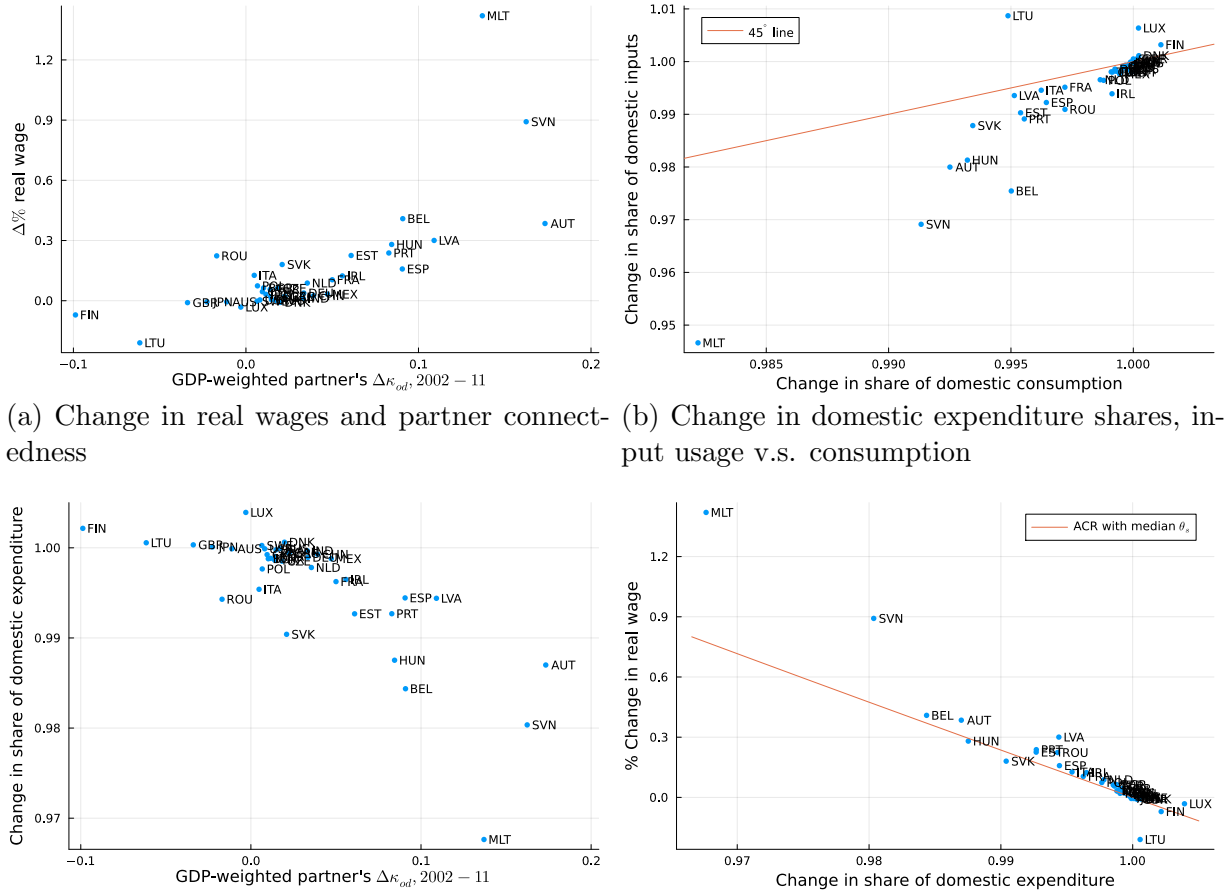
Since we have solved $\hat{R}_{os}$ as a function of $\{\hat{w}_o\}$, we can substitute them into equation (C-6) and obtain N non-linear equations. We iterate to solve $\hat{w}_o$ using a fixed-point algorithm. Intuitively, if the left-hand side of (C-6) is higher than the right-hand side, labor demand falls short of labor supply, and we reduce the guess of $\hat{w}_o$. We raise the guess of $\hat{w}_o$ vice versa.

D Additional Quantitative Results

Figure D-1 examines the relationships between the changes in connectedness, domestic expenditure shares and real wages in the 2002–2011 counterfactual of Section 6. Panel (a) plots the change in real wages against the GDP-weighted average change in the partners' connectedness. The changes in domestic expenditure shares in final consumption and in intermediate inputs are not the same, but they are highly correlated, as illustrated by Panel (b); Panel

(c) plots the change in the overall domestic expenditure share against the GDP-weighted average change in the partners' connectedness. In the original Eaton-Kortum model with a single-sector, round-about economy, there is a log-linear relationship between the change in real wages and the domestic expenditure share, i.e., $\hat{w}/\hat{P} = (\widehat{s_{dom}})^{-\frac{1}{\theta a_L}}$ (see also Arkolakis, Costinot and Rodríguez-Clare (2012)). In our model with multiple sectors and input-output linkages, we do not have such a tight relationship. When we compute $(\widehat{s_{dom}})^{-\frac{1}{\theta a_L}}$ in each country using the overall change in domestic expenditure share and the median trade elasticity θ, we find that it slightly under-predicts the welfare changes in each country, as illustrated by the red solid line in Panel (d); the actual changes in $\hat{w}/\hat{P}$ are nevertheless highly correlated with $(\widehat{s_{dom}})^{-\frac{1}{\theta a_L}}$.

Figure D-1: Impact of the change in air connectedness from 2002 to 2011



(a) Change in real wages and partner connectedness

(b) Change in domestic expenditure shares, input usage v.s. consumption

(c) Change in domestic expenditure shares and partner connectedness

(d) Change in real wages and domestic expenditure shares

Notes. This figure examines the relationship between different variables concerning the impact of the change in air connectedness from 2002 to 2011. In Panels (a) and (c), the change in partner connectedness averages $\Delta\kappa_{od}$ across partner countries, weighting each partner by its base-period GDP. The red solid line in Panel (d) is the welfare changes predicted by the formula $(\widehat{s_{dom}})^{-\frac{1}{\theta a_L}}$, where $\widehat{s_{dom}}$ is the change in overall domestic expenditure share and θ is the median trade elasticity across sectors.

Table D-1 benchmarks the 2002–2011 counterfactual against the observed changes in the WIOD over the same window, country by country, as discussed in Section 6. The first three columns compare trade shares: for each country, the foreign share of total absorption in 2002, its observed change between 2002 and 2011, and the change predicted by the counterfactual. The last two columns compare welfare changes. The benchmark in the fourth column is the welfare change implied by the observed changes in trade shares under the assumption that changes in sectoral trade costs are their only driver. It is computed with the sufficient-statistics formula of Costinot and Rodríguez-Clare (2014) (the perfect-competition case of their equation 28),

$$\ln \frac{\hat{w}_d}{\hat{P}_d} = -\,\xi_d' \left(I - A_d\right)^{-1} \left(\frac{\ln \hat{\pi}_{dds}}{\theta_s} \right)_{s=1}^{S}, \tag{D-1}$$

where $\hat{\pi}_{dds}$ is the observed change in country d's domestic trade share in sector s between the two periods, ξ_d is the vector of consumption shares, and $(A_d)_{sk} = a_{dk,s}$ collects the input-output coefficients, so that the weights $\xi_d'(I - A_d)^{-1}$ measure the exposure of consumption to each sector's price through the input-output linkages. The last column reports the welfare change predicted by the counterfactual, as in Table 9.[37]

For the world as a whole (the last row of the table, in which the share columns are ratios of world aggregates and the welfare columns are GDP-weighted averages), the foreign share of absorption was 10.3% in 2002 and rose by 2.0 percentage points by 2011, of which the counterfactual predicts 0.10 percentage points, or 4.9%. The observed sectoral shifts imply a world welfare gain of 1.49% through formula (D-1), of which the counterfactual predicts 0.03%, or 2.1%. The shares are considerably larger for the countries with the largest improvements in connectedness: in Austria, the counterfactual accounts for 1.00 of the observed 2.07 percentage-point increase in the foreign share (48%) and 0.38 of the implied 0.94% welfare gain (41%); in Belgium, the corresponding shares are 44% and 14%.

[37] The country-level benchmark can differ in sign from the change in the aggregate foreign share. The aggregate share weights sectors by absorption, whereas formula (D-1) adds up the log changes in sectoral domestic shares divided by the sectoral trade elasticities, so it is dominated by the sectors with the lowest elasticities—machinery (1.45) and rubber and plastics (1.67), and to a lesser extent food and chemicals—in which even a modest change in a small domestic share is a large log change. The largest benchmark gains, in Latvia, Estonia and Malta, come from the collapse of their domestic shares in these sectors, although their aggregate foreign shares fell; the negative benchmarks for Canada and Russia reflect rising domestic shares in the same sectors.

Table D-1: Benchmarking the 2002–2011 counterfactual against observed changes

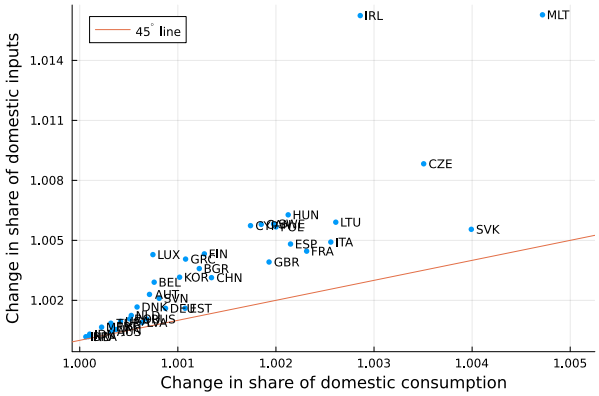
Country	$1 - s_{dom}^{2002}$	$\Delta(1 - s_{dom})^{data}$	$\Delta(1 - s_{dom})^{model}$	$\Delta\%\,(w/P)^{CRC}$	$\Delta\%\,(w/P)^{model}$
Australia	10.41	0.36	0.009	1.17	-0.01
Austria	22.97	2.07	1.002	0.94	0.38
Belgium	29.46	2.52	1.103	3.02	0.41
Bulgaria	24.95	-2.25	0.091	-0.96	0.06
Brazil	7.02	0.45	0.021	0.21	0.01
Canada	19.20	-3.03	0.039	-3.92	0.02
China	8.25	-0.10	0.066	-0.94	0.03
Cyprus	25.21	-4.65	0.057	-1.67	0.04
Czechia	23.72	5.34	0.113	8.28	0.06
Germany	16.06	4.02	0.089	3.66	0.04
Denmark	21.27	2.39	-0.049	3.31	-0.01
Spain	14.05	0.48	0.478	1.54	0.16
Estonia	29.60	-5.51	0.516	12.19	0.23
Finland	16.38	2.97	-0.181	1.36	-0.07
France	12.97	1.61	0.326	1.96	0.10
United Kingdom	13.44	2.70	-0.029	4.56	-0.01
Greece	17.55	-0.85	0.052	0.75	0.02
Hungary	28.27	6.81	0.895	6.49	0.28
Indonesia	12.83	-0.85	0.102	1.43	0.03
India	7.63	2.70	0.030	1.55	0.01
Ireland	35.74	2.93	0.226	2.77	0.12
Italy	11.83	2.60	0.406	1.93	0.13
Japan	4.92	2.74	-0.009	0.70	-0.01
South Korea	13.73	5.89	0.083	2.11	0.02
Lithuania	27.33	-0.16	-0.041	5.00	-0.21
Luxembourg	48.92	1.85	-0.201	8.01	-0.03
Latvia	23.11	-3.57	0.431	12.51	0.30
Mexico	14.22	3.30	0.103	2.35	0.03
Malta	37.94	-4.31	2.008	12.02	1.42
Netherlands	23.51	4.98	0.167	2.29	0.09
Poland	15.50	5.86	0.198	4.92	0.07
Portugal	17.80	-0.76	0.601	0.27	0.24
Romania	19.12	-1.57	0.461	-1.04	0.22
Russia	12.23	-0.69	0.006	-1.25	0.00
Slovakia	29.42	-1.02	0.677	8.78	0.18
Slovenia	25.58	0.18	1.462	8.30	0.89
Sweden	19.21	1.93	-0.021	1.41	-0.00
Turkey	11.07	4.54	0.123	0.35	0.06
USA	6.63	2.08	0.010	1.35	0.00
World	10.31	1.98	0.097	1.49	0.03

Notes. This table compares the 2002–2011 connectedness counterfactual with the observed changes in the WIOD over the same window. $1 - s_{dom}^{2002}$ is the foreign share of total absorption in 2002, in percent. $\Delta(1 - s_{dom})^{data}$ and $\Delta(1 - s_{dom})^{model}$ are the observed and the model-predicted changes in the foreign share from 2002 to 2011, in percentage points. $\Delta\%\,(w/P)^{CRC}$ is the welfare change implied by the observed sectoral domestic-share changes through the sufficient-statistics formula, the perfect-competition case of eq. (28) in Costinot and Rodríguez-Clare (2014); $\Delta\%\,(w/P)^{model}$ is the change in real wages predicted by the counterfactual, as in Table 9. In the last row, the share columns are ratios of world aggregates, and the welfare columns are weighted averages of all countries using the base-period GDP as the weights.

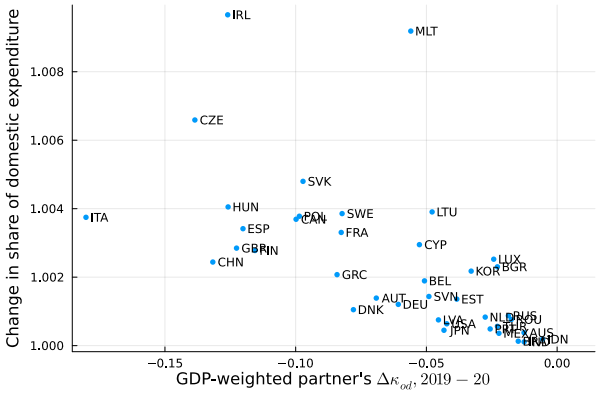
Figure D-2 repeats the plots of Figure D-1 for the drop in air connectedness from 2019 to

2020, and Table D-2 reports the country-level results of that counterfactual. Panel (a) plots the change in the domestic expenditure share in intermediate inputs against the change in the share in final consumption. The two are highly correlated, and every country lies above the 45-degree line: input sourcing responds more strongly than final consumption, consistent with our mechanism operating through the reliability of imported inputs. Panel (b) shows that the increases in the overall domestic expenditure share line up with the declines in partner connectedness: the countries whose connectedness with partners fell the most reallocate the most expenditure toward domestic goods. Panel (c) plots the change in real wages against the change in the domestic expenditure share. As in the 2002–2011 exercise, the one-sector round-about benchmark with the median trade elasticity (the red solid line) tracks the model closely: real wages fall where domestic expenditure shares rise. The panel that plots real wages against the change in partner connectedness is omitted, since it appears as Panel (b) of Figure 4 in the main text.

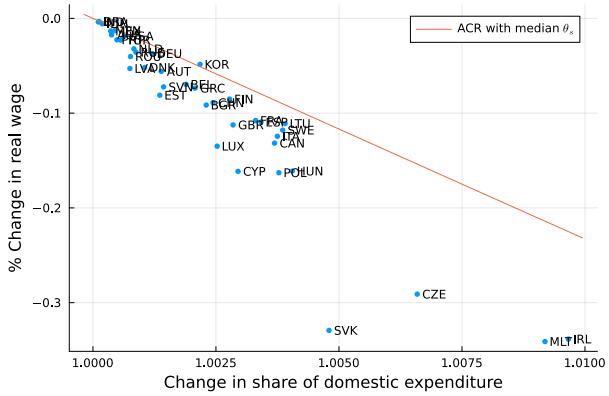
Figure D-2: Impact of the change in air connectedness from 2019 to 2020



(a) Change in domestic expenditure shares, input usage v.s. consumption



(b) Change in domestic expenditure shares and partner connectedness



(c) Change in real wages and domestic expenditure shares

Notes. This figure examines the relationship between different variables concerning the impact of the change in air connectedness from 2019 to 2020. In Panel (b), the change in partner connectedness averages $\Delta\kappa_{od}$ across partner countries, weighting each partner by its base-period GDP. The red solid line in Panel (c) is the welfare changes predicted by the formula $(\widehat{s_{dom}})^{-\frac{1}{\theta a_L}}$, where $\widehat{s_{dom}}$ is the change in overall domestic expenditure share and θ is the median trade elasticity across sectors.

Table D-2: Impact of the change in air connectedness from 2019 to 2020

Country	κ_o^{2019}	$\Delta\kappa_o$	$\Delta\% \, s_{dom}^{final}$	$\Delta\% \, s_{dom}^{input}$	$\Delta\% \, s_{dom}$	$\Delta\% \, w/P$
Australia	0.13	-0.01	0.03	0.04	0.04	-0.02
Austria	0.46	-0.07	0.07	0.23	0.14	-0.06
Belgium	0.53	-0.05	0.08	0.29	0.19	-0.07
Bulgaria	0.26	-0.02	0.12	0.36	0.23	-0.09
Brazil	0.14	-0.01	0.01	0.02	0.01	-0.00
Canada	0.49	-0.10	0.19	0.58	0.37	-0.13
China	0.30	-0.13	0.13	0.31	0.24	-0.09
Cyprus	0.21	-0.05	0.17	0.57	0.29	-0.16
Czechia	0.36	-0.14	0.35	0.88	0.66	-0.29
Germany	0.71	-0.06	0.09	0.16	0.12	-0.04
Denmark	0.57	-0.08	0.06	0.17	0.11	-0.05
Spain	0.47	-0.12	0.21	0.48	0.34	-0.11
Estonia	0.20	-0.04	0.11	0.16	0.14	-0.08
Finland	0.51	-0.12	0.13	0.43	0.28	-0.09
France	0.53	-0.08	0.23	0.45	0.33	-0.11
United Kingdom	0.77	-0.12	0.19	0.39	0.28	-0.11
Greece	0.39	-0.08	0.11	0.41	0.21	-0.07
Hungary	0.41	-0.13	0.21	0.63	0.41	-0.16
Indonesia	0.06	-0.01	0.01	0.03	0.02	-0.01
India	0.11	-0.01	0.01	0.02	0.01	-0.00
Ireland	0.54	-0.13	0.29	1.62	0.97	-0.34
Italy	0.52	-0.18	0.26	0.49	0.37	-0.12
Japan	0.44	-0.04	0.04	0.05	0.04	-0.01
South Korea	0.46	-0.03	0.10	0.32	0.22	-0.05
Lithuania	0.26	-0.05	0.26	0.59	0.39	-0.11
Luxembourg	0.27	-0.02	0.07	0.43	0.25	-0.13
Latvia	0.28	-0.05	0.06	0.09	0.08	-0.05
Mexico	0.35	-0.02	0.02	0.07	0.04	-0.01
Malta	0.30	-0.06	0.47	1.63	0.92	-0.34
Netherlands	0.73	-0.03	0.05	0.12	0.08	-0.03
Poland	0.42	-0.10	0.20	0.57	0.38	-0.16
Portugal	0.45	-0.03	0.03	0.07	0.05	-0.02
Romania	0.29	-0.02	0.05	0.11	0.08	-0.04
Russia	0.11	-0.02	0.07	0.11	0.09	-0.04
Slovakia	0.17	-0.10	0.40	0.56	0.48	-0.33
Slovenia	0.14	-0.05	0.08	0.21	0.14	-0.07
Sweden	0.30	-0.08	0.20	0.58	0.39	-0.12
Turkey	0.25	-0.02	0.03	0.09	0.06	-0.02
USA	0.42	-0.04	0.04	0.09	0.06	-0.02
Weighted Average	0.45	-0.06				-0.05

Notes. This table reports the impact of the change in air connectedness from 2019 to 2020. κ_o^{2019} and $\Delta\kappa_o$ are the level in 2019 and the change from 2019 to 2020 of country o's connectedness with its partners, averaged across partners using their base-period GDP as the weights. s_{dom}^{final}, s_{dom}^{input} and s_{dom} denote the domestic expenditure share in final consumption, intermediate input usage and total absorption. w/P indicates real wage, in which the price index is calculated as $P_d = \prod_s P_{ds}^{\xi_{ds}}$. The last row reports the weighted average of all countries, using the base-period GDP as the weights.

Table D-3 reports the group-level reallocation of Russia's goods trade underlying Figure 5 in Section 6.4.

Table D-3: Reallocation of Russia's foreign goods trade by partner group

Partner group	Imp%$_r$	Imp%$_{cf}$	ΔImp(pp)	Exp%$_r$	Exp%$_{cf}$	ΔExp(pp)
Domestic (share of total)	66.68	67.21	0.54	76.57	76.77	0.20
Still-connected	17.91	18.47	0.56	2.62	2.66	0.04
W. Europe	45.83	45.01	-0.81	10.11	8.69	-1.41
E. Europe & Baltics	7.85	7.76	-0.09	2.18	2.00	-0.18
East Asia	4.48	4.46	-0.03	3.17	3.12	-0.05
North America	5.21	5.28	0.07	2.26	2.26	-0.00
Rest of World	18.73	19.02	0.29	79.67	81.27	1.60
Total	100.00	100.00	-0.00	100.00	100.00	-0.00

Notes. Goods trade only (WIOD sectors 1–16). The first row reports Russia's domestic share of its total goods absorption (import columns) and of its total goods sales (export columns), including the domestic flow in the denominator: this is the reshoring margin discussed in the text. In the group rows, shares are computed over foreign partners only, excluding Russia's domestic flow, so they describe the reallocation of the foreign trade basket rather than the reshoring margin; Δ is in percentage points. Groups are defined in the notes to Figure 5.

D.1 Sensitivity of the Counterfactuals to the Trade Elasticity θ_s

Our baseline counterfactuals use the sectoral trade elasticities θ_s from Caliendo and Parro (2015). Table D-4 recomputes the changes in real wages from the improvement in air connectedness from 2002 to 2011 under two alternative parameter sets: the elasticities estimated by Giri et al. (2020), aggregated to the WIOD sectors, and a common elasticity equal to the median of the Caliendo and Parro (2015) sectoral values.[38] The cross-country pattern of welfare changes is very similar across the three sets of elasticities.

[38]Sectors not covered by Giri et al. (2020) are assigned the median of the covered sectors, mirroring the treatment of non-tradable sectors in the baseline (see footnote 9 for a comparison of the two sets of estimates). The elasticity of substitution ρ_s is from Soderbery (2015) throughout.

Table D-4: Sensitivity to θ_s: change in air connectedness from 2002 to 2011

	$\Delta\% \, w/P$		
Country	CP θ_s	GYY θ_s	median θ
AUS	-0.01	-0.00	-0.00
AUT	0.38	0.35	0.37
BEL	0.41	0.33	0.36
BGR	0.06	0.06	0.06
BRA	0.01	0.01	0.01
CAN	0.02	0.02	0.02
CHN	0.03	0.02	0.03
CYP	0.04	0.04	0.04
CZE	0.06	0.06	0.06
DEU	0.04	0.04	0.04
DNK	-0.01	-0.01	-0.01
ESP	0.16	0.15	0.16
EST	0.23	0.23	0.25
FIN	-0.07	-0.07	-0.07
FRA	0.10	0.10	0.11
GBR	-0.01	-0.01	-0.01
GRC	0.02	0.02	0.02
HUN	0.28	0.27	0.29
IDN	0.03	0.02	0.03
IND	0.01	0.01	0.01
IRL	0.12	0.11	0.12
ITA	0.13	0.13	0.13
JPN	-0.01	-0.01	-0.01
KOR	0.02	0.02	0.02
LTU	-0.21	-0.21	-0.21
LUX	-0.03	-0.02	-0.03
LVA	0.30	0.28	0.29
MEX	0.03	0.03	0.03
MLT	1.42	1.23	1.34
NLD	0.09	0.06	0.06
POL	0.07	0.07	0.07
PRT	0.24	0.23	0.23
ROU	0.22	0.19	0.19
RUS	0.00	0.01	0.01
SVK	0.18	0.12	0.13
SVN	0.89	0.85	0.90
SWE	-0.00	-0.00	-0.00
TUR	0.06	0.06	0.06
USA	0.00	0.00	0.00
Weighted Average	0.03	0.03	0.03

Notes. This table reports the percentage change in real wages from the change in air connectedness from 2002 to 2011 (the counterfactual in Table 9) under three sets of trade elasticities: the baseline θ_s from Caliendo and Parro (2015) (CP), the θ_s from Giri et al. (2020) (GYY), and a common θ equal to the median of the CP sectoral values. The last row reports the weighted average of all countries, using the base-period GDP as the weights.

D.2 Excluding Air-Intensive Sectors from the Counterfactual

For goods intensively shipped by air, part of the effect of air connectedness on trade could in principle operate through air freight costs rather than communication costs. Our estimation-stage evidence indicates that this is not what drives our results: when we exclude products typically transported by air, the estimates are, if anything, larger (Online Appendix B.4). We therefore view the communication-cost mechanism as operating in every sector, and our baseline counterfactuals accordingly shock all sectors. As a deliberately conservative robustness check, here we recompute the 2002–2011 counterfactual excluding the air-intensive sectors from the shock entirely.

We classify a goods sector as air-intensive if its air share of US import value—the value of imports shipped by air divided by the total import value, aggregated from HS 6-digit products to WIOD sectors—is above a cutoff. This is the same value-based share that underlies the product exclusions in our estimation-stage robustness checks. Table D-5 reports the share for each of the sixteen goods sectors: with a 20% cutoff, three goods sectors are air-intensive, and with a 10% cutoff, seven of the sixteen.[39] The air shares are unavailable for service sectors; since the delivery of traded services relies heavily on air travel, we treat every service sector as air-intensive whenever a cutoff is applied. We then recompute the counterfactual holding communication costs in the air-intensive sectors fixed at their base-year levels, so that the change in κ_{od} affects only the goods sectors below the cutoff.

[39] At the 20% cutoff: chemicals and chemical products; electrical and optical equipment; and manufacturing, not elsewhere classified and recycling. The 10% cutoff adds textiles and textile products; leather and footwear; basic metals and fabricated metal; and machinery, not elsewhere classified.

Table D-5: Air share of US import value by goods sector

Sector	Air share of import value (%)
Agriculture, hunting, forestry and fishing	9.6
Mining and quarrying	0.8
Food, beverages and tobacco	2.7
Textiles and textile products	13.0
Leather, leather products and footwear	13.7
Wood and products of wood and cork	0.7
Pulp, paper, printing and publishing	5.8
Coke, refined petroleum and nuclear fuel	1.2
Chemicals and chemical products	48.1
Rubber and plastics	6.5
Other non-metallic mineral	7.0
Basic metals and fabricated metal	17.1
Machinery, not elsewhere classified	15.5
Electrical and optical equipment	50.0
Transport equipment	7.9
Manufacturing, not elsewhere classified; recycling	31.2

Notes. This table reports, for each of the sixteen WIOD goods sectors, the share of US import value shipped by air, in percent: the value of US imports shipped by air divided by the total import value, aggregated from HS 6-digit products to WIOD sectors using import values as weights.

Table D-6 reports the changes in real wages. Excluding the air-intensive sectors mechanically reduces the magnitudes, because a large share of trade no longer responds to the shock. Even in the most stringent variant—holding communication costs fixed in all service sectors and in the seven goods sectors with an air share above 10%—about a third of the GDP-weighted average welfare gain from the 2002–11 improvement in air connectedness remains (0.01% versus 0.03% in the baseline), and two thirds remain under the 20% cutoff (0.02%). The cross-country pattern is essentially unchanged under both cutoffs. However, we believe that the communication-cost mechanism operates in every sector, including the air-intensive ones, and we still shock all sectors with $\Delta\kappa_{od}$ in our baseline.

Table D-6: Excluding air-intensive sectors: change in air connectedness from 2002 to 2011

	$\Delta\% \, w/P$		
Country	baseline	excl. air share $\geq$ 20%	excl. air share $\geq$ 10%
AUS	-0.01	-0.01	-0.02
AUT	0.38	0.22	0.15
BEL	0.41	0.26	0.20
BGR	0.06	0.03	0.02
BRA	0.01	0.00	0.00
CAN	0.02	0.01	0.01
CHN	0.03	0.01	0.00
CYP	0.04	0.05	0.05
CZE	0.06	0.04	0.03
DEU	0.04	0.02	0.01
DNK	-0.01	-0.00	-0.00
ESP	0.16	0.10	0.07
EST	0.23	0.11	0.08
FIN	-0.07	-0.04	-0.03
FRA	0.10	0.06	0.04
GBR	-0.01	-0.00	-0.01
GRC	0.02	0.01	0.00
HUN	0.28	0.12	0.07
IDN	0.03	0.02	0.02
IND	0.01	0.01	0.00
IRL	0.12	0.02	0.02
ITA	0.13	0.08	0.05
JPN	-0.01	-0.00	-0.00
KOR	0.02	0.01	0.01
LTU	-0.21	-0.13	-0.09
LUX	-0.03	0.01	0.00
LVA	0.30	0.18	0.12
MEX	0.03	0.02	0.01
MLT	1.42	0.56	0.40
NLD	0.09	0.05	0.05
POL	0.07	0.04	0.03
PRT	0.24	0.15	0.10
ROU	0.22	0.17	0.07
RUS	0.00	-0.00	-0.01
SVK	0.18	0.10	0.09
SVN	0.89	0.57	0.36
SWE	-0.00	0.00	-0.00
TUR	0.06	0.03	0.01
USA	0.00	0.00	0.00
Weighted Average	0.03	0.02	0.01

Notes. This table reports the percentage change in real wages from the change in air connectedness from 2002 to 2011 (the counterfactual in Table 9) when air-intensive sectors keep their base-year communication costs. The baseline column shocks all sectors; in the cutoff columns, goods sectors with an air share of US import value at or above 20% or 10% and all service sectors keep their base-year communication costs. The air share of a goods sector is the value of US imports shipped by air divided by the total import value, aggregated from HS 6-digit products, as reported in Table D-5. The last row reports the weighted average of all countries, using the base-period GDP as the weights.

D.3 Sensitivity of the Counterfactuals to the Treatment of Trade Imbalances

Our counterfactuals model trade imbalances as exogenous lump-sum transfers D_d and, in the baseline, scale them with income in the counterfactual equilibrium, $D'_d = \hat{w}_d D_d$, so that total expenditure changes in proportion to the wage, $\hat{E}_d = \hat{w}_d$, and the change in the real wage coincides with the change in real absorption. The natural alternative, also discussed by Costinot and Rodríguez-Clare (2014), is to hold the transfers fixed in numeraire units, $D'_d = D_d$. Table D-7 recomputes the welfare effects of the change in air connectedness from 2002 to 2011 under the two conventions. The choice is immaterial for our results: the welfare changes differ by less than 0.02 percentage points in every country, because the connectedness shock is small relative to the terms-of-trade adjustments that separate the two conventions in large-tariff experiments.

Table D-7: Sensitivity to the treatment of trade imbalances: change in air connectedness from 2002 to 2011

	$\Delta\% \, w/P$	
Country	$D'_d = \hat{w}_d D_d$	$D'_d = D_d$
AUS	-0.01	-0.01
AUT	0.38	0.39
BEL	0.41	0.41
BGR	0.06	0.06
BRA	0.01	0.01
CAN	0.02	0.02
CHN	0.03	0.03
CYP	0.04	0.04
CZE	0.06	0.06
DEU	0.04	0.04
DNK	-0.01	-0.01
ESP	0.16	0.16
EST	0.23	0.22
FIN	-0.07	-0.07
FRA	0.10	0.10
GBR	-0.01	-0.01
GRC	0.02	0.02
HUN	0.28	0.28
IDN	0.03	0.03
IND	0.01	0.01
IRL	0.12	0.13
ITA	0.13	0.13
JPN	-0.01	-0.01
KOR	0.02	0.02
LTU	-0.21	-0.21
LUX	-0.03	-0.03
LVA	0.30	0.30
MEX	0.03	0.03
MLT	1.42	1.41
NLD	0.09	0.09
POL	0.07	0.07
PRT	0.24	0.24
ROU	0.22	0.22
RUS	0.00	0.00
SVK	0.18	0.18
SVN	0.89	0.89
SWE	-0.00	-0.00
TUR	0.06	0.06
USA	0.00	0.00
Weighted Average	0.03	0.03

Notes. This table reports the percentage change in real wages from the change in air connectedness from 2002 to 2011 under two conventions for the exogenous transfers D_d in the counterfactual equilibrium: the baseline, in which the transfers scale with income ($D'_d = \hat{w}_d D_d$), and the alternative, in which they are held fixed in numeraire units ($D'_d = D_d$). The last row reports the weighted average of all countries, using the base-period GDP as the weights.

E Detour Distances and Counterfactual Connectedness

This appendix provides details on the construction of the counterfactual distances and the counterfactual connectedness index used in Section 6.4.

E.1 Finding the Detour Path

We discretize the Earth's surface into a $1° \times 1°$ grid. Each cell is either *free*, meaning aircraft may traverse it, or *blocked*, meaning it lies within the no-fly zone, to which we add a 0.5° buffer. We represent the grid as a graph whose nodes are the centres of free cells and whose edges connect each cell to its (up to) eight neighbours, horizontally, vertically and diagonally. The weight assigned to each edge is the great-circle distance between the two adjacent cell centres, so summing edge weights along a path gives the total flight distance accumulated by traversing that sequence of cells.

The standard algorithm for finding the shortest path on such a graph is Dijkstra's algorithm, which Besedeš et al. (2024) use in a related setting. Dijkstra explores nodes in order of their accumulated cost $g(n)$ from the origin, at each step expanding whichever unvisited node is cheapest to reach so far, until the destination is reached. It is guaranteed to find the globally optimal path, but it does so blindly: it expands nodes in all directions from the origin with no regard for where the destination lies, and can waste considerable effort on nodes far from any promising route.

We instead use the A* algorithm, a directed generalization of Dijkstra that exploits knowledge of the destination's location to focus the search. At each step, A* selects the node that minimizes

$$f(n) = g(n) + h(n),$$

where $g(n)$ is the accumulated great-circle cost from j to n along the best path found so far and $h(n)$ is a heuristic estimate of the remaining cost from n to the destination i. Nodes that are both cheap to reach and close to the destination are prioritized, so the search concentrates along promising corridors rather than exploring the entire graph. We set

$$h(n) = d_{gc}(n,\, i),$$

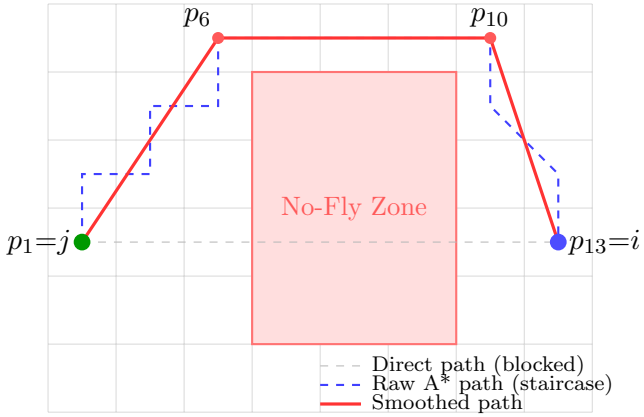
the direct great-circle distance from n to i. Because the great-circle arc is the shortest possible path between two points on a sphere, $h(n)$ never overestimates the true remaining cost. This property, called admissibility, guarantees that A* returns the globally optimal path, the same answer Dijkstra would give, while typically expanding far fewer nodes. The saving is what makes the exercise feasible: more than four million potential routes must first

be screened for intersection with the no-fly zone, and more than 1.3 million of them then require a full shortest-path search.

E.2 The Staircase Problem and Path Smoothing

Because the graph has only 1° resolution and movement is restricted to eight directions, the raw A* path zig-zags between grid steps even when the true optimal route would be a smooth geodesic, which overstates the distance. Figure E-1 illustrates the artifact on a schematic grid. The raw path is symmetric by construction: the sequence of cells visited from j to i is the reverse of the sequence from i to j.[40]

Figure E-1: A* on a discrete grid: staircase artifact and path smoothing



Notes. The grid represents the 1° × 1° discretization. The shaded rectangle is the no-fly zone. The direct great-circle path (grey dashed) is blocked by the obstacle. The raw A* path (blue dashed) bypasses the obstacle via 12 alternating grid steps ($p_1, \ldots, p_{13}$), producing the characteristic staircase pattern. The smoothed path (red solid) replaces these 12 steps with three direct arcs, $p_1 \to p_6$, $p_6 \to p_{10}$ and $p_{10} \to p_{13}$, each of which avoids the no-fly zone.

Suppose the raw A* path traverses n waypoints $p_1 = j, p_2, \ldots, p_n = i$. We smooth it by the following greedy shortcutting procedure applied from p_1:

1. Test whether the direct great-circle arc $p_1 \to p_{n-1}$ avoids the no-fly zone. If it does, the smoothed path is $p_1 \to p_{n-1} \to p_n$.

2. If $p_1 \to p_{n-1}$ is blocked, test $p_1 \to p_{n-2}$. If this arc is clear, determine the best tail from p_{n-2} to p_n: if the direct arc $p_{n-2} \to p_n$ is also clear, the path is $p_1 \to p_{n-2} \to p_n$; otherwise it is $p_1 \to p_{n-2} \to p_{n-1} \to p_n$.

3. More generally, we find the largest $k < n$ for which the arc $p_1 \to p_k$ is clear, and recursively apply the same procedure from p_k to p_n.

[40] A finer grid resolution would reduce the staircase artifact, but given the number of routes that must be screened and searched, a 1° resolution combined with smoothing is a computational necessity.

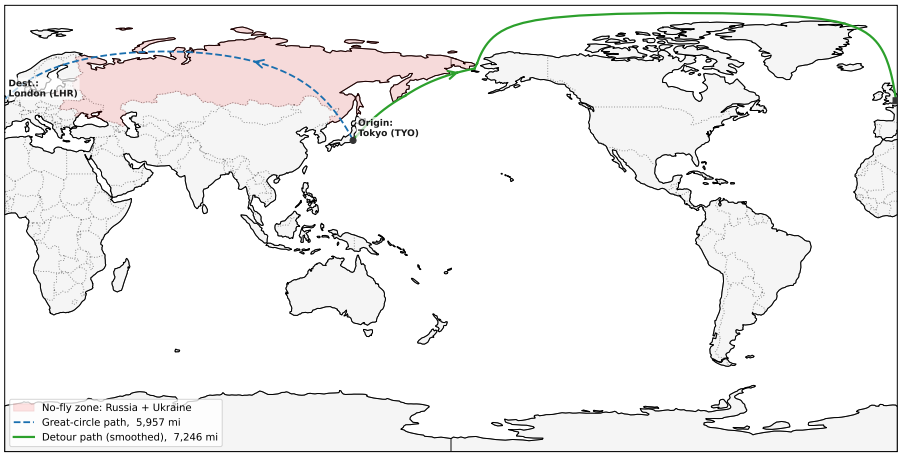
This yields a path composed of a small number of direct great-circle arcs whose total length is never greater than the raw A* distance. In the example of Figure E-1, the procedure identifies p_6 as the farthest waypoint directly reachable from p_1, then p_{10} as the farthest reachable from p_6, and finally connects p_{10} directly to $p_{13} = i$, reducing 12 staircase steps to three arcs.

Because the shortcutting procedure is applied in a fixed direction, the smoothed path from j to i may differ slightly from the smoothed path from i to j, and their lengths can differ marginally. We therefore compute both directions and retain the shorter smoothed distance as d_{ji}^{detour}, consistent with our goal of identifying the minimum feasible detour. The difference is negligible in practice: on average the longer direction exceeds the shorter one by 0.25%.

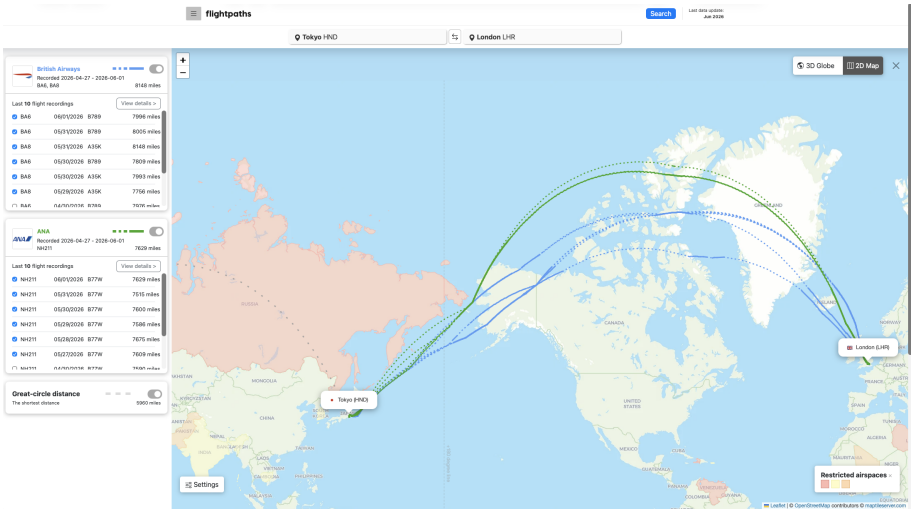
E.3 Validation and the Distribution of Detour Distances

Figure E-2 validates the procedure against observed flight paths on the London–Tokyo route, one of the corridors most affected by the closure. Panel (a) is the path our algorithm returns: the great-circle path crosses Russia, and the shortest feasible path routes north over the Arctic and then south around the closed airspace, which raises the distance from roughly 5,900 to more than 7,000 miles and moves the pair from below to above the 6,000-mile threshold. Panel (b) shows the routes British Airways and All Nippon Airways actually flew after February 2022, which follow the same corridor.

Figure E-2: The London–Tokyo route after the closure of Russian and Ukrainian airspace



(a) Theoretical optimal detour path

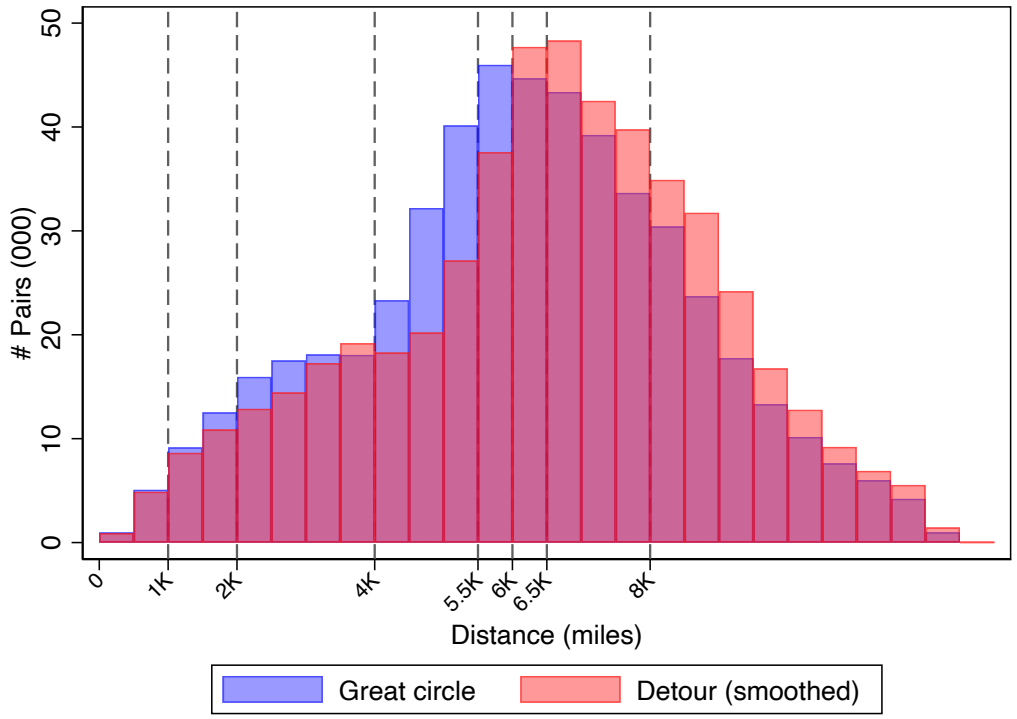


(b) Real detour path

Notes. Panel (a) shows the theoretical route identified by our algorithm: the blue path is the great-circle path between London and Tokyo, the green path is the shortest path that avoids Russian and Ukrainian airspace after greedy shortcutting, and the shaded region is the no-fly zone. Panel (b) shows the routes actually flown by British Airways and All Nippon Airways after February 2022, obtained from https://www.flightpaths.com. Both route over the Arctic and the North Atlantic, as the algorithm predicts.

Figure E-3 overlays the distributions of great-circle and smoothed detour distances across all affected grid pairs. The two are similar at short and very long ranges, but the detour distribution shifts right in the middle of the range: routes that would have overflown Russian or Ukrainian airspace are pushed into longer distance bins, and many cross from below to above 6,000 miles. This is the variation the counterfactual exploits.

Figure E-3: Great-circle versus detour distance distribution



Notes. The figure plots the distribution of great-circle distances and smoothed detour distances across all grid pairs. Each bin covers a 500-mile interval. Grid pairs whose origin or destination lies within the no-fly zone (with a 0.5° buffer) are excluded from the detour distribution, as their detour distance is infinite.

E.4 From Distance to Counterfactual Connectedness

Table E-1 reports the regressions we use to translate the change in distance into a change in connectedness. The dependent variable is a dummy equal to one if grids j and i are connected by a weekly direct flight, and the regressors are indicators for great-circle distance groups, with $(6\text{K}, 6.5\text{K}]$ miles omitted. Column (1) includes no fixed effects and is the specification we use for the counterfactual, so that the predicted connectedness of a grid pair depends only on its distance; the remaining columns add origin and destination grid fixed effects and country-pair fixed effects, and show that the distance profile is stable across them. The estimates decline steeply and monotonically with distance up to the omitted group and are negative beyond it. The profile is steepest at short distances, falling from 0.0356 in the $(0, 1\text{K}]$ bin to 0.0138 in $(1\text{K}, 2\text{K}]$ and 0.0021 in $(2\text{K}, 4\text{K}]$, which is why rerouted short-haul pairs contribute materially to the aggregate change even though they never approach the 6,000-mile threshold.

Table E-1: Predicting grid-pair connectedness using distance groups

Dep. Var.	$\mathbb{1}(\text{Has Direct Flight}_{ji})$			
	(1)	(2)	(3)	(4)
$\mathbb{1}\left(d_{ji}^{GC} \in (0, 1K] \text{ miles}\right)$	0.0356^{a}	0.0332^{a}	0.0275^{a}	0.0253^{a}
	(0.0004)	(0.0004)	(0.0005)	(0.0005)
$\mathbb{1}\left(d_{ji}^{GC} \in (1K, 2K] \text{ miles}\right)$	0.0138^{a}	0.0138^{a}	0.0120^{a}	0.0114^{a}
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
$\mathbb{1}\left(d_{ji}^{GC} \in (2K, 4K] \text{ miles}\right)$	0.0021^{a}	0.0021^{a}	0.0028^{a}	0.0028^{a}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$\mathbb{1}\left(d_{ji}^{GC} \in (4K, 5.5K] \text{ miles}\right)$	0.0008^{a}	0.0001^{c}	0.0009^{a}	0.0006^{a}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$\mathbb{1}\left(d_{ji}^{GC} \in (5.5K, 6K] \text{ miles}\right)$	0.0006^{a}	-0.0002^{a}	0.0005^{a}	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$\mathbb{1}\left(d_{ji}^{GC} \in (6.5K, 8K] \text{ miles}\right)$	-0.0001^{c}	0.0000	-0.0003^{a}	-0.0003^{a}
	(0.0000)	(0.0001)	(0.0001)	(0.0001)
$\mathbb{1}\left(d_{ji}^{GC} > 8K \text{ miles}\right)$	-0.0005^{a}	-0.0009^{a}	-0.0008^{a}	-0.0009^{a}
	(0.0000)	(0.0001)	(0.0001)	(0.0001)
Origin Grid FE	N	Y	N	Y
Destination Grid FE	N	Y	N	Y
Origin-Destination-Country-Pair FE	N	N	Y	Y
without FE	Y	N	N	N
Observations	4745706	4745706	4743636	4743636
# of Grid Pairs	4745706	4745706	4743636	4743636
Within R-squared	0.016	0.013	0.003	0.002
R-squared	0.016	0.076	0.059	0.111

Notes. The sample consists of all the airport-grid pairs in our dataset, with the connection dummy measured in 2019. The omitted distance group is $d_{ji}^{GC} \in (6\text{K}, 6.5\text{K}]$ miles, so all coefficients measure connectedness relative to that group. Column (1) is the specification used to construct the counterfactual connectedness index. Standard errors clustered at the grid-pair level in parentheses. Significance levels: a: 0.01, b: 0.05, c: 0.10.

Evaluating Column (1) at the great-circle and at the detour distance and differencing gives, for each airport-grid pair (j', i'), the predicted change in the probability of a direct connection,

$$\Delta\text{Prob}_{j'i'} = \hat{\mathbb{1}}(\text{HasDirectFlight}_{j'i'}^{detour}) - \hat{\mathbb{1}}(\text{HasDirectFlight}_{j'i'}^{gc}),$$

where the first term is set to zero for grid pairs whose detour distance is infinite. Since grid-level connectedness is defined through the 200-mile neighbourhoods of Figure 1, we form $\overline{\Delta\text{Prob}_{j'i'}}$, the average of $\Delta\text{Prob}_{j'i'}$ over the airport-grid pairs within those neighbourhoods, and set

$$\Delta\kappa_{ji} = \overline{\Delta\text{Prob}_{j'i'}},$$

so that the regression supplies the change in connectedness directly and no counterfactual level of the index needs to be constructed. We apply this neighbourhood average to every

pair of grids with an airport, so that the shock also reaches routes that were close to the margin of being connected. Since equation (8) aggregates grid pairs linearly, the grid-level changes aggregate to the country-pair changes $\Delta\kappa_{od}$ that enter the counterfactual. Across the 1,482 WIOD country pairs, the index falls on average by 0.008, about 1.8% of its 2019 level.