Communication Costs, Direct Flights and International Trade*

Taiji Furusawa Chang Sun Heiwai Tang Jiaxu Zhang †

Aug, 2024

Abstract

We build a general equilibrium model of endogenous communication, quality control and trade. We derive a structural gravity equation from the model and show that exogenous communication costs raise the costs of quality control and have a larger impact on products with a lower elasticity of substitution. In our empirical application, we estimate the impact of direct flight connectedness on communication costs using the gravity equation. We overcome the identification challenge using an instrumental variable constructed based on the discontinuity of direct flights at around 6,000-mile distance due to air travel regulations. We find that air connectedness increases trade, especially for products with a low elasticity. We combine the empirical estimates and our equilibrium model to quantify the aggregate impact of air connectedness.

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

*We thank Bingjing Li, Haishi (Harry) Li and participants at the Asia-Pacific Trade Seminars for their helpful comments. Financial support from HKGRF (project code: 17503221) 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 critical for understanding the patterns of trade and the welfare gains from trade, and economists have long been interested in measuring such costs. As Anderson and Van Wincoop (2004) point out, the trade costs that are needed to match the aggregate trade data are very large (a tax equivalent of 170%), while the typical "trade costs" such as transportation and tariff barriers are very low. (Hummels, 1999; Harrigan, 1993) One potential explanation for the "black matters" in the 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 in trade costs.

Incorporating endogenous communication and quality control in Furusawa, Inui, Ito and Tang (2017) into Eaton and Kortum (2002), we 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 by conducting costly communication with the seller. On the other hand, exogenous communication costs, such as the difficulty of business travel, increase the total communication costs for any given level of 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 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.

Direct flights are a great medium to enhance business travel and face-to-face communication. In our empirical analysis, we focus on the air connectedness between countries due to direct flights as a key factor for 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 $1° \times 1°$ grid cells. For any pair of grids in two countries, we treat 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 the 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 the trade flows. However, an Ordinary Least Square (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 around 6,000-mile distance to construct the IV. The regulations in the United States and Europe restrict hours of operation for flight crew members within a 24-hour period and lead to additional crew in any flight of more than 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. Though the overall connectedness between two countries may be heavily affected by their distance or the share of grid pairs that are out of this window, the event that grid pairs have a distance below 6,000 miles conditional on being in the [5500, 6500] window is plausibly random. 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 using both OLS and IV regressions. We find positive and statistically significant coefficients of the interaction term mentioned above, 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 4.2% increase in trade when air connectedness increases by ten percentage points.

We show that our estimates are robust to a battery of robustness checks. Our main 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, is still 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, the historical trade flows strongly predicts current trade flows, but they do not diminish our IV estimates of direct flight connectedness. We address other concerns and show that our main results are robust to excluding products that are mainly shipped by air, to using alternative trade elasticities when constructing the composite elasticity, to modified versions of the IV and to different ways to construct air connectedness.

In the last section of the paper, 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 2016, all countries in our sample except for Australia experienced an improvement in air connectedness with the other countries. Correspondingly, our model predicts that these countries will have higher real wages due to the improvement in air connectedness. The model reveals that small countries with large improvements in air connectedness, such as Malta, Ireland, and Slovenia, have the largest gains in real wages, ranging from 1.56% to 0.88%.

In 2020, the COVID-19 pandemic caused disruptions to international travel, and all countries in our sample experienced a decline in air connectedness from 2019 to 2020. We perform another counterfactual by changing the bilateral air connectedness from the level of 2019 to that of 2020. Similar to the first counterfactual, we find that small countries with large declines in air connectedness, such as Malta, Ireland, Slovakia and the Czech Republic, had the largest declines in real wages, ranging from 0.35% to 0.29%. A caveat when interpreting the quantitative results is that we only identify a particular mechanism that improvement or worsening of air connectedness affects trade and the overall effect of air connectedness may be larger due to other mechanisms.

Our paper makes contributions to three strands of literature. First, it helps us better understand the barriers to doing business between sellers and producers. Previous research has highlighted different types of barriers that are not typically considered in the gravity literature (Anderson and Van Wincoop, 2004). Existing studies have focused on search and contractual frictions. For example, Bernard, Moxnes and Saito (2018) build a model in which business travel can reduce search frictions and improve outsourcing efficiency, and they provide evidence for the mechanism using the opening of a high-speed train line in Japan. Startz (2021) incorporates both search frictions and contractual frictions in a model of traders endogenously determining whether to travel to buy or order goods from distant suppliers, and uses novel survey data of Nigerian traders to support and quantify the model. We focus on quality control and communication costs, a mechanism that has not been studied much before.[1] One advantage of our approach is that we obtain structural gravity equations and can estimate the impact of shocks on communication costs using aggregate trade data.

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

Second, we contribute to the literature on the impact of air travel on international business.[2] 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 efforts 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 bearings on the major predictions of the Eaton and Kortum (2002) model other than affecting the level of the price index in all countries, but it plays a crucial role in our model since 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 novel work by Söderlund (2023), who uses the liberalization of the Soviet Airspace as a shock to causally identify the impact of direct flights on international trade, though his approach is largely reduced-form.

Third, this paper also contributes to understanding the impact of regulations on travel costs and other outcomes. For example, restrictions on air space 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 have higher growth during twenty years. 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.

The paper proceeds as follows. Section 2 introduces our theoretical model and derives the structural gravity equation that connects communication costs and trade flows. In Section 3, we describe the data and how we construct the key variables in our empirical and quantitative

[2] 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).

analysis. We introduce the identification strategy and regression specifications in Section 4 and report the estimated effect of direct flights on communication costs in Section 5. Section 6 closes the model in general equilibrium and conducts counterfactual analysis. We conclude in Section 7.

2 Model

In this section, we build a general equilibrium model of sourcing with endogenous communication efforts. Our model combines key elements of Eaton and Kortum (2002) and Furusawa et al. (2017) and delivers a structural gravity equation that connects communication costs and 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 $\mathcal{S}$. In each destination d and sector s, final goods producers combine intermediate goods from various locations and sell their output to consumers. We assume perfect competition and common technologies among these producers, so they price at marginal costs. 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. 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 has a linear production function using only labor as input and draws its productivity $z_{os}(\omega)$ to produce any variety from a Fréchet distribution with a cumulative distribution function $G_{os}(z) = \exp(-T_{os} z^{-\theta_s})$, 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 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, q, 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 from the integration in equation (1). To choose q, the buyer has to engage in communication with individual suppliers, which raises the unit cost of the inputs by a multiple of $e^{m_{od}q}$. We refer to m_{od} as the exogenous communication costs. 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 assume that buyers in destination d and sector s choose a common q for all suppliers from origin o, which we denote as q_{ods}.[3]

We first derive the unit cost for the composite input bundle in sector s, given the unit cost and the probability of meeting the buyer's standard of each variety. Specifically, the buyer chooses quantities $x(\omega)$ to the expenditure on inputs:

$$P_{ds} \equiv \min_{x(\omega)} \int_0^1 p(\omega)x(\omega)d\omega \quad \text{s.t.} \quad \int_0^1 \left(q_{o(\omega),d,s}\right) x(\omega)^{1-1/\rho_s} d\omega \geq 1,$$

where $p(\omega)$ denotes the unit cost of intermediate inputs including communication costs but excluding attrition from input failures, and $o(\omega)$ denotes the country from which the buyer sources variety ω. Note that we have applied the assumption that q is common for all suppliers from origin o when writing the production function. The optimization problem delivers the unit cost

$$P_{ds} = \left[\int_0^1 \left(\left(q_{o(\omega),d,s}\right)^{-\tilde{\rho}_s} p(\omega)\right)^{1-\rho_s} d\omega\right]^{\frac{1}{1-\rho_s}}, \quad \tilde{\rho}_s \equiv \frac{\rho_s}{\rho_s - 1}. \tag{2}$$

Since $q_{o(\omega),d,s} < 1$, the price $p(\omega)$ is adjusted by the factor $\left(q_{o(\omega),d,s}\right)^{-\tilde{\rho}_s} > 1$, which reflects the cost of product failure. Given the same $q_{o(\omega),d,s}$, the factor is larger when inputs are less substitutable with each other, i.e., a smaller ρ_s and a larger $\tilde{\rho}_s \equiv \frac{\rho_s}{\rho_s-1}$.

As in the original EK model, buyers source input ω from the lowest-cost location, taking into account the cost of product failures:

$$\min_o \left\{ q_{ods}^{-\tilde{\rho}_s} e^{m_{od}q_{ods}} \frac{c_{os}t_{od}}{z_{os}(\omega)} \right\}. \tag{3}$$

From this equation, it is clear that the problem of choosing q_{ods} can be separated from the sourcing problem. Regardless of ω, production costs and the iceberg trade costs, the buyer

[3]We can relax this assumption by assuming that suppliers from one origin country are divided into several mutually exclusive and collectively exhaustive sets,each with a positive mass, and that the buyer chooses a common q within each set. We need the sets to contain a positive mass of suppliers to apply the law of large numbers within each set. The optimal q, however, does not depend on the distribution of productivities within each set and must be the same across all suppliers in an o-d pair. This can be seen from the adjusted unit cost that enters P_{ds} in equation (2): $q(\omega)^{-\tilde{\rho}_s}p(\omega) = q(\omega)^{-\tilde{\rho}_s}e^{m_{od}q(\omega)}\frac{c_{os}t_{od}}{z_{os}(\omega)}$. Minimizing this term, we obtain $q(\omega) = \tilde{\rho}_s/m_{od}$, which is independent of $z_{os}(\omega)$.

minimizes $q_{ods}^{-\tilde{\rho}_s} e^{m_o d q_{ods}}$ and obtains

$$q_{ods} = \frac{\tilde{\rho}_s}{m_{od}}, \quad q_{ods}^{-\tilde{\rho}_s} e^{m_{od} q_{ods}} = \left(\frac{e m_{od}}{\tilde{\rho}_s}\right)^{\tilde{\rho}_s}.$$

On the other hand, under the restriction that $\theta_s > \rho_s - 1$, we can apply the properties of the Fréchet distribution and obtain

$$P_{ds} = \gamma \left[\sum_o T_{os} \left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od}\right)^{-\theta_s}\right]^{-1/\theta_s}, \tag{4}$$

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

It is immediate that the parameter ρ_s plays an important role in our model: higher communication costs 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, 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

$$\log X_{ods} = \delta_{ds} + \log T_{os} - \theta_s \log c_{os} - \theta_s \tilde{\rho}_s \log m_{od} - \theta_s \log t_{od}, \tag{5}$$

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 the 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 identifying the impact of shocks, such as the establishment of direct flights, on communication costs. In Sections 4 and 5, we use a regression specification based on equation (5) and estimate the impact of direct flights.

As another advantage of our model, we can apply the "exact had 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.

3 Data

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

3.1 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 also make our sample size more manageable, we select the top 100 countries in total imports, accounting for more than 95% of world trade.

3.2 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 information about individual flight stages of international scheduled services from 1989 to 2020[4]. 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 positive average seats available.[5]

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 city pairs 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 (4 million pairs), and we find around 7,000 city pairs with weekly direct flights.

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

[5]TFS provides data aggregated at origin-destination-carrier-aircraft-type level. We do not know the number of seats available on every flight, but we know the average seats available within a carrier and aircraft-type cell.

3.3 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.[6] We obtain the population of each grid from the Geographically based Economic (G-Econ 4.0) data.[7] 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 use Google Map Geocoding API to obtain the longitude and latitude of the cities in TFS data and assign each city to a $1^\circ \times 1^\circ$ grid based on its coordinates.[8] We calculate the connection index between countries o and d, κ_{od}, using the following formula:

$$\kappa_{od} = \sum_{j \in o} \omega_j \sum_{i \in d} \omega_i \kappa_{ji}, \tag{6}$$

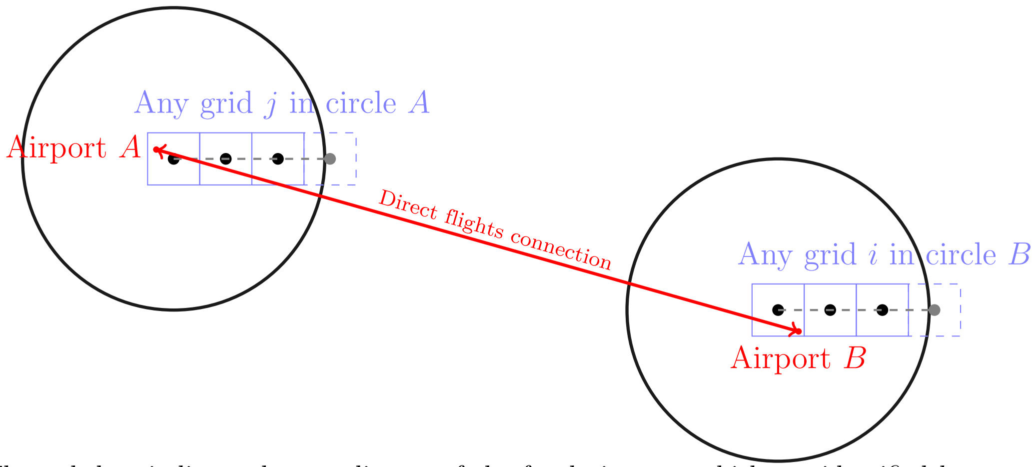
where we abuse the notation slightly and use o and d also to denote the set of grids in the two countries. j denotes a grid in the origin country, and i denotes a grid in the destination country. κ_{ji} is a dummy variable, which equals one if grid centers of j and i are within a 200-mile radius of two 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 the population in 1990, the earliest year in which G-Econ reports population data, to calculate the weights. By construction, κ_{od} equals κ_{do}. 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.

[6] In practice, the area of each grid depends on its location. One latitude is equivalent to 69 miles anywhere on the earth. At the equator, one longitude is also equivalent to 69 miles. But the distance of one longitude gradually shortens when moving from the equator to the poles.

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

[8] Since Google Map API may return multiple coordinates for one TFS city, we obtain unique coordinates of each TFS airport city using two extra steps. First, we use each country's shape file and remove coordinates outside the corresponding country of each city. Second, for cities with multiple coordinates, if the maximum distance between these coordinates is below 100 miles, we use the simple average of these coordinates as the coordinate of the city. If the maximum distance is above 100 miles, we manually search 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 as connected if the two airports have established a direct flight connection.

4 Empirical Specifications

This section describes the empirical method we use and the rationale for using it. 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 flight on communication costs.

4.1 Estimation Equation

We use the structural gravity equation (5) to estimate how direct flights affect the communication costs, m_{od}. We parameterize the communication costs as a function of the connection index in equation (6):

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

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 is aligns well with our approach to measuring air connectedness and captures rich topology features of the origin and destination countries.

Substituting the expressions of m_{od} and t_{od} in the gravity equation (5), we obtain our estimation equation

$$\log X_{ods} = \beta \kappa_{od} \times \theta_s \tilde{\rho}_s - \sum_{n=1}^{6} \gamma_n \lambda_{od,n} \times \theta_s + \delta_{ds} + \delta_{os} + \epsilon_{ods}, \tag{7}$$

where δ_{ds} and δ_{os} are destination-sector and origin-sector fixed effects, and ϵ_{ods} is the error term.

In our main specification, we treat every HS 6-digit product as a "sector". 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.

[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 a sectoral Eaton-Kortum model.

4.2 Identification using Instrumental Variables

The simplest approach to estimate equation (7) is to apply Ordinary Least Square (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 hours of operation for flight crew members within a 24-hour period and lead to additional crew in any flight of more than 12 hours, which correspond to approximately 6,000 miles of distance. Combined with the arrival of new aircraft that could afford such long-haul flights in 1989, later years saw a lot more increase in the number of flights in 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 A, 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 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. Mathematically, our IV is

$$z_{od} = \sum_{j \in o, i \in d} \omega_{ji} \mathbb{1}\left(d_{ji} \leq 6K | d_{ji} \in (5.5K, 6.5K]\right) \tag{8}$$

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 $\omega_{ji} \equiv \frac{\text{pop}_j \times \text{pop}_i}{\sum_{j \in o, i \in d} \text{pop}_j \times \text{pop}_i \times \mathbb{1}(d_{ji} \in (5.5K, 6.5K])}$.[10] 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,

[10] The weights here are slightly different from those in equation (6). 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 $\omega_j \times \omega_i = \frac{\text{pop}_j \times \text{pop}_i}{\sum_{j \in o, i \in d} \text{pop}_j \times \text{pop}_i}$.

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

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

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 fewer than 25% of the country pairs are connected by weekly direct flights. The instrumental variable, z_{od}, has an average of 54.4%. 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 below 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, conditional on being in the 5,500 and 6,500-mile 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	66064.675	322668.333	6	48	340	1344.5	5929	29036.5	122101	904428	7752241
# Grid Pairs in (5.5K,6.5K)	3024	11761.403	52447.329	1	2	40	210	1142.5	5901.5	19779	184183	978547

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 6.6% and that between the USA and Russia is 6.4%, 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: We write smaller countries in terms of area as "Country 1" in the table. We rank rows of the table by the value of the connection index. Our constructed index is symmetric and independent of the ordering of Country 1 and Country 2.

In Table 3, we show that the IV z_{od} 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 a 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) or grid pairs that are much more likely to be connected (less than 5,500 miles apart). However, we see the conditional probability z_{od} has a statistically significant effect on the 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 (7). 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)
$\sum_{ij\in od} w_{ij}\mathbb{1}\left(d_{ij}\leq 6K \vert d_{ij}\in(5.5K,6.5K]\text{ miles}\right)$	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 country-origin country level. Significance levels: a: 0.01, b: 0.05, c: 0.10.

5 Empirical Results

In this section, we present our main empirical results. We show that both OLS and IV regressions suggest that direct flights reduce communication costs as predicted by our model. We also provide a battery of robustness checks.

5.1 Baseline Results

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 (7) using the full sample, including all origin-destination-product level trade flows among the 100 major trading partners that account for more than 95% of global trade. 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 (lower 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 4.2% increase in trade when air connectedness increases by ten percentage points.[11]

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 a 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 country-origin country level. Significance levels: a: 0.01, b: 0.05, c: 0.10.

5.2 Robustness Checks

In this section, we conduct several robustness checks. These tests demonstrate that our results remain significant after selecting different industry-level parameters, controlling for other factors that may affect trade, constructing instrumental variables using alternative

[11] 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.

parameters, and adjusting the way the connection indices are calculated.

5.2.1 Control Trade Volumes in Earlier Years

Our instrumental variable approach relies on the exclusion restriction that, conditional on control variables, other factors that influence trade (ϵ_{ods} in equation 7), are orthogonal to the instrument $z_{od} \times \theta_s \tilde{\rho}_s$. If this assumption does not hold, we may expect that ϵ_{ods} affects both trade and the endogenous variable of interest instrumented by $z_{od} \times \theta_s \tilde{\rho}_s$. In this section, we control for historical trade that captures the persistent components in ϵ_{ods} that may be correlated with our instrument. In addition, it may also improve our control of the impact of conventional trade costs.

In Table 5, we run the same regressions as in Table 4 but control for trade flows $X_{ods,0}$ in earlier years. In the first three columns, we control for trade flows in 1989, the earliest year for which we have trade data and also the year when new long-haul aircraft were just introduced and the discontinuous drop in direct flights at around 6,000 miles had not emerged yet (see more evidence in Appendix A). One issue with the earlier trade data is that they contain a lot more zeros than in later years, partly because we have more countries due to the dissolution of the Soviet Union. To avoid losing too many observations, we transform $X_{ods,0}$ using the inverse-hyperbolic-sine transformation, i.e., $\log\left(X_{ods,0} + \sqrt{X_{ods,0}^2 + 1}\right)$ and use the transformed variable as the control. We see that the trade flows in 1989 positively predict the trade flows in 2016 in both OLS and IV regressions. They are also positively associated with the overall air connectedness in 2016, suggesting an endogeneity problem with OLS regressions (see Column 3). However, we find the IV estimates similar to our main regressions. In columns 4-6, we control trade flows in 1992 instead of 1989, and the estimates are almost the same.

Table 5: Robustness Checks: Trade in Earlier Years as Controls

	Sample in (5.5K,6.5K)					
	OLS	IV	First Stage	OLS	IV	First Stage
	(1)	(2)	(3)	(4)	(5)	(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)	
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K \mid d_{ij} \in (5.5K, 6.5K]\right)$			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 country-origin country 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 origin country-destination country-HS6 sector level. To prevent losing too many observations due to zero historical trade flows, we do an inverse-hyperbolic-sine transformation to the trade flows in 1989 and 1992.

5.2.2 Exclude Products Transported by Air

In this section, we change our sample products to exclude 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 [12] to calculate the share of air shipping of each product. In our original sample, about half of the products have an air transport share below 5%. In Table 6, we focus on products that are not typically shipped by air, i.e., with an air transport share below 5% or 1% and re-run our main specifications. The coefficients are positive and significant at conventional levels.

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

Table 6: Robustness Checks: Exclude Products Transported by Air

	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)	
$\theta_s\tilde{\rho}_s \times \sum_{ij\in od} w_{ij}\mathbb{1}\left(d_{ij} \leq 6K \mid d_{ij} \in (5.5K, 6.5K]\right)$			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 country-origin country 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.

5.2.3 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, but they also cover fewer industries and lead to a smaller sample.

Table 7 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 slightly higher than those using θ_s from Caliendo and Parro (2015).

Table 7: 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)	
$\theta_s\tilde{\rho}_s \times \sum_{ji\in od} w_{ji}\mathbb{1}\left(d_{ij} \leq 6K \mid d_{ij} \in (5.5K, 6.5K]\right)$				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 country-origin country 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 8: 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)	
$\tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K \mid d_{ij} \in (5.5K, 6.5K]\right)$				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 country-origin country 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.

5.2.4 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 9 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-statistics become smaller when we use a narrower window to construct the IV. It is straightforward that when the window becomes narrower, it is harder to predict the overall air connectedness between two countries, especially for large countries that may have many grid pairs that are outside the window. We find that the IV would not significantly predict overall connectedness if we further narrowed the window.

Table 9: 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)
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K \vert d_{ij} \in (5.5K, 6.5K]\right)$	0.091^a		
	(0.015)		
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K \vert d_{ij} \in (5.6K, 6.4K]\right)$		0.075^a	
		(0.019)	
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K \vert d_{ij} \in (5.7K, 6.3K]\right)$			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 country-origin country 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 changing the window when constructing the IVs, we also try to improve the first stage 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 10, 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 10: 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)
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K \mid d_{ij} \in (5.5K, 6.5K]\right)$	0.091^a	0.064^a	0.072^a
	(0.015)	(0.024)	(0.022)
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K \mid d_{ij} \in (5.5K, 6.5K]\right)$ $\times$ % in (2K,5.5K)		0.053^b	
		(0.025)	
$\theta_s \tilde{\rho}_s \times \sum_{ij \in od} w_{ij} \mathbb{1}\left(d_{ij} \leq 6K \mid d_{ij} \in (5.5K, 6.5K]\right)$ $\times$ 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 country-origin country level. Significance levels: a: 0.01, b: 0.05, c: 0.10.

5.2.5 Alternative Air Connectedness κ_{od}

In this section, we show that our results are robust to alternative ways to construct the connection index κ_{od}. In our baseline, we treat two grids as connected if both of them fall within a 200-mile radius of airports that have weekly direct flights in between. 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 11, 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 we increase the radius. One reason may be that the connected airports do not affect grids that are too far away. When we see 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 that are within 100 miles, the mean connection index is only 0.031, and intuitively, the average connection index between country pairs increases in radius selection since countries are more likely to be connected if airports can affect a wider range of cities.

Table 11: 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 $\text{Index}_{od} \times \theta_s \tilde{\rho}_s$	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)
$\theta_s \tilde{\rho}_s \times \sum_{ji \in od} w_{ji} \mathbb{1}\left(d_{ij} \leq 6K \mid d_{ij} \in (5.5K, 6.5K]\right)$	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	0.104	0.116	0.127	0.146
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 country-origin country level. Significance levels: a: 0.01, b: 0.05, c: 0.10. θ_s is from Caliendo and Parro (2015).

In Table 12, 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 significant and positive, and the estimates are also larger when we increase the threshold. This suggests that more frequent direct flights reduce communication costs further and lead to more trade flows.

Table 12: 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)	
$\theta_s \tilde{\rho}_s \times \sum_{ji \in od} w_{ji} \mathbb{1}\left(d_{ij} \leq 6K \mid d_{ij} \in (5.5K, 6.5K]\right)$			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 country-origin country 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. Twice-weekly (daily) connection is defined as there are more than 104 (365) round-trip direct flights between two countries in a year.

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 the improvement in air connectedness from 2002 to 2016 and the quick drop from 2019 to 2020 due to the Covid-19 pandemic.

6.1 General Equilibrium and Hat Algebra

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 flight. 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 (4)) depends on the unit cost of the input bundle, c_{os}. The latter can be written as

$$c_{os} = \Upsilon_{os} w_o^{a_{L,os}} \prod_k P_{ok}^{a_{ok,s}}. \tag{9}$$

$a_{L,os}$ 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

$$R_{os} = \sum_{d} \pi_{ods} \left(\xi_{ds} E_d + \sum_{k} a_{ds,k} R_{dk} \right), \tag{10}$$

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

$$\pi_{ods} \equiv \frac{X_{ods}}{X_{ds}} = \frac{T_{os} \left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od} \right)^{-\theta_s}}{\sum_o T_{os} \left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od} \right)^{-\theta_s}}. \tag{11}$$

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 (10) 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 (10) given shares, prices, and the exogenous transfers.

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

$$w_o L_o = \sum_{s} a_{L,os} R_{os}. \tag{12}$$

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

$$\frac{m'_{od}}{m_{od}} = e^{-\beta \left(\kappa'_{od} - \kappa_{od} \right)}, \tag{13}$$

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 (4) in changes:

$$\hat{P}_{ds}^{-\theta_s} = \sum_{o} \pi_{ods} \left(\hat{m}_{od}^{\tilde{\rho}_s} \hat{c}_{os} \right)^{-\theta_s}. \tag{14}$$

The change in the cost of the input bundle, $\hat{c}_{os}$, can be obtained from equation (9):

$$\hat{c}_{os} = \hat{w}_o^{a_{L,os}} \prod_k \hat{P}_{ok}^{a_{ok,s}}. \tag{15}$$

Combining $2N \times S$ nonlinear equations (14) and (15), 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

$$\hat{\pi}_{ods} = \left[\frac{T_{os}\left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od}\right)^{-\theta_s}}{\sum_o T_{os}\left(m_{od}^{\tilde{\rho}_s} c_{os} t_{od}\right)^{-\theta_s}} \right]^{-1} \frac{T_{os}\left((m'_{od})^{\tilde{\rho}_s} c'_{os} t_{od}\right)^{-\theta_s}}{\sum_o T_{os}\left((m'_{od})^{\tilde{\rho}_s} c'_{os} t_{od}\right)^{-\theta_s}} = \frac{\left(\hat{m}_{od}^{\tilde{\rho}_s} \hat{c}_{os}\right)^{-\theta_s}}{\sum_o \pi_{ods}\left(\hat{m}_{od}^{\tilde{\rho}_s} \hat{c}_{os}\right)^{-\theta_s}} \tag{16}$$

The goods market clearing conditions (10) can be written in changes as

$$\hat{R}_{os} = \sum_d \frac{\pi_{ods} \xi_{ds} E_d}{R_{os}} \hat{\pi}_{ods} \hat{E}_d + \sum_{d,k} \frac{\pi_{ods} a_{os,k} R_{dk}}{R_{os}} \hat{\pi}_{ods} \hat{R}_{dk}, \tag{17}$$

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, which we set to be the same as the base-period values, D_d, in our application. Similar to equation (10), equation (17) is a linear system of $\hat{R}_{os}$ and we can apply matrix inversion to obtain the solution.

The labor market clearing conditions (12) can be written in changes as

$$\hat{w}_o = \sum_s \frac{a_{L,os} R_{os}}{w_o L_o} \hat{R}_{os}. \tag{18}$$

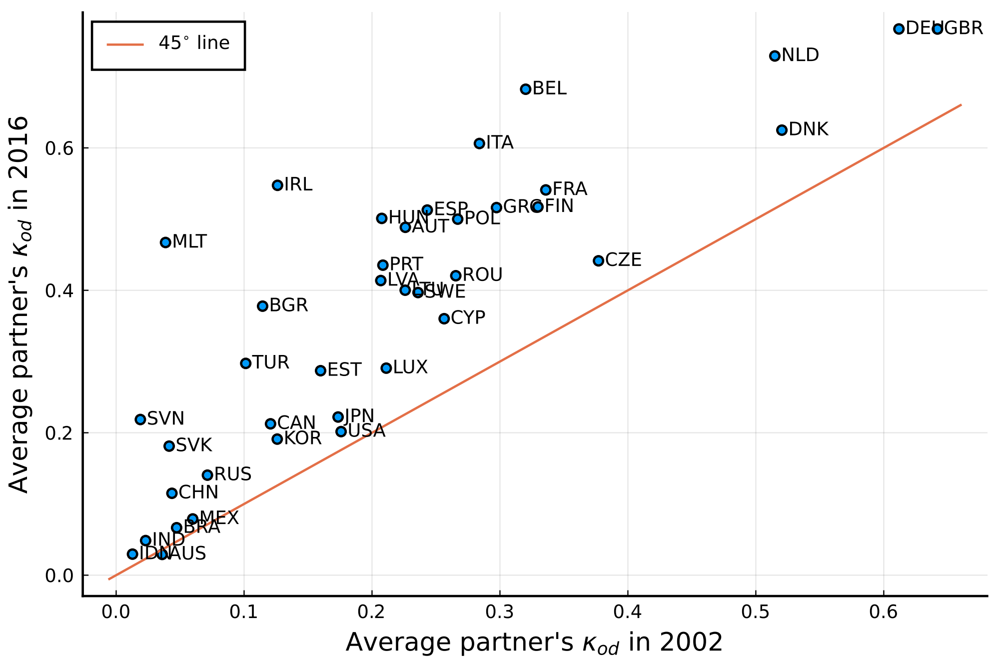
Since we have solved $\hat{R}_{os}$ as a function of $\{\hat{w}_o\}$, we can substitute them into equation (18) and obtain N non-linear equations. We iterate to solve $\hat{w}_o$ using a fixed-point algorithm. Intuitively, if the left-hand side of (18) is higher than the right-hand side, labor demand falls short of labor supply, and we reduce the guess of $\hat{w}_o$. We raise the guess of $\hat{w}_o$ vice versa.

6.2 Improvement in Air Connectedness from 2002 to 2016

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. We focus on 39 countries with non-missing data on bilateral air connectedness κ_{od}. The set of countries are listed in Table 13.

Figure 2 compares air connectedness in 2016 with that in 2002. For each country o, we calculate the average of κ_{od} across all of its partners, d, in the base year 2002 and the later year 2016. In almost all countries, connectedness with foreign countries improved during this period. European countries, such as Malta, Ireland, Italy and Belgium, saw increases in κ_{od} above 0.30. Meanwhile, the air connectedness of the United States only increased from 0.18 to 0.20. Using our empirical estimate $\beta = 0.026$, we use equation (13) 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, ρ_s.[13]

Figure 2: Average κ_{od} in 2016 v.s. 2002



In the last four columns of Table 13, we present percentage changes in equilibrium outcomes due to the change in air connectedness from 2002 to 2016. We report the change in domestic expenditure shares in final consumption, intermediate input usage and total absorption, and the change in real wages. The change in domestic expenditure shares in final consumption and intermediate input are not the same, but they are highly correlated as illustrated by Panel (b) of Figure 3. Drops in these shares mean that countries spend more

[13]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.

on foreign goods and that consumer welfare improves.[14] This is confirmed by the change in real wages in the last column. Consistent with the changes in air connectedness, Malta and Ireland have the highest gains from the improvement in air connectedness during this period (1.56% and 1.00%, respectively). Despite not having the highest improvement in air connectedness, Slovenia ranks third in the increase of real wages (0.88%). This is because the country started from very low levels of κ_{od} in 2002 and is also a very small economy. As in typical trade models, gains more from trading with foreign countries are larger for smaller economies.

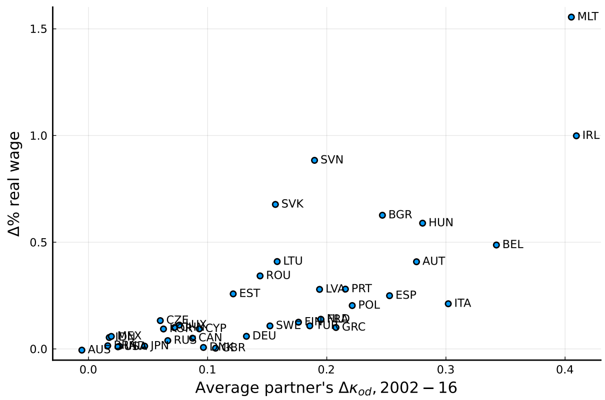
[14]In the original EK model with a single-sector, round-about economy, there is a log-linear relationship between the change in real wages and domestic expenditure share, i.e., $\hat{w}/\hat{P} = (\widehat{s_{dom}})^{-1/\theta a_L}$. (see also Arkolakis, Costinot and Rodríguez-Clare (2012)) In our model with multiple sectors and input-output linkages, we do not have such a tight relationship. When we compute $(\widehat{s_{dom}})^{-1/\theta a_L}$ in each country using the overall change in domestic expenditure share (Column 5 in Table 13) 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) of Figure 3. But the actual changes in $\hat{w}/\hat{P}$ are highly correlated with $(\widehat{s_{dom}})^{-1/\theta a_L}$.

Table 13: Impact of the change in air connectedness from 2002 to 2016

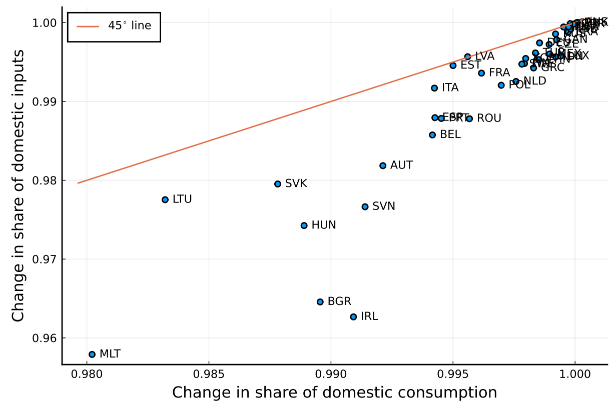
Country	$\kappa_{od,2002}$	$\Delta\kappa_{od}$	$\Delta\%\, s_{dom}^{final}$	$\Delta\%\, s_{dom}^{input}$	$\Delta\%\, s_{dom}$	$\Delta\%\, w/P$
AUS	0.04	-0.01	0.01	0.00	0.01	-0.01
AUT	0.26	0.28	-0.79	-1.82	-1.23	0.41
BEL	0.33	0.34	-0.58	-1.42	-1.01	0.49
BGR	0.12	0.25	-1.04	-3.54	-2.21	0.63
BRA	0.06	0.02	-0.03	-0.12	-0.07	0.01
CAN	0.13	0.09	-0.08	-0.22	-0.14	0.05
CHN	0.05	0.07	-0.15	-0.47	-0.35	0.10
CYP	0.27	0.09	-0.02	-0.01	-0.02	0.09
CZE	0.44	0.06	-0.11	-0.28	-0.20	0.13
DEU	0.64	0.13	-0.15	-0.26	-0.20	0.06
DNK	0.53	0.10	0.01	-0.00	0.00	0.01
ESP	0.25	0.25	-0.57	-1.21	-0.87	0.25
EST	0.16	0.12	-0.50	-0.54	-0.52	0.26
FIN	0.34	0.18	-0.21	-0.52	-0.36	0.13
FRA	0.37	0.20	-0.38	-0.64	-0.50	0.14
GBR	0.66	0.11	-0.01	-0.02	-0.01	0.00
GRC	0.30	0.21	-0.17	-0.57	-0.30	0.10
HUN	0.22	0.28	-1.11	-2.57	-1.78	0.59
IDN	0.01	0.02	-0.08	-0.43	-0.25	0.05
IND	0.03	0.03	-0.02	-0.08	-0.05	0.01
IRL	0.14	0.41	-0.91	-3.73	-2.26	1.00
ITA	0.29	0.30	-0.58	-0.83	-0.70	0.21
JPN	0.18	0.05	-0.05	-0.05	-0.05	0.01
KOR	0.13	0.06	-0.20	-0.45	-0.34	0.09
LTU	0.23	0.16	-1.68	-2.25	-1.88	0.41
LUX	0.22	0.08	-0.06	-0.43	-0.24	0.11
LVA	0.22	0.19	-0.44	-0.43	-0.44	0.28
MEX	0.07	0.02	-0.10	-0.40	-0.21	0.06
MLT	0.04	0.41	-1.98	-4.21	-2.81	1.56
NLD	0.53	0.19	-0.24	-0.75	-0.47	0.14
POL	0.29	0.22	-0.30	-0.79	-0.54	0.20
PRT	0.21	0.22	-0.55	-1.22	-0.84	0.28
ROU	0.28	0.14	-0.43	-1.22	-0.79	0.34
RUS	0.07	0.07	-0.08	-0.14	-0.11	0.04
SVK	0.05	0.16	-1.22	-2.05	-1.66	0.68
SVN	0.02	0.19	-0.86	-2.34	-1.57	0.88
SWE	0.25	0.15	-0.22	-0.53	-0.36	0.11
TUR	0.10	0.19	-0.16	-0.38	-0.26	0.11
USA	0.18	0.02	-0.03	-0.05	-0.04	0.01
Weighted Average	0.24	0.08				0.06

Notes: This table reports the impact of the change in air connectedness from 2002 to 2016. $s_{dom}^{final}, s_{dom}^{input}$ 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.

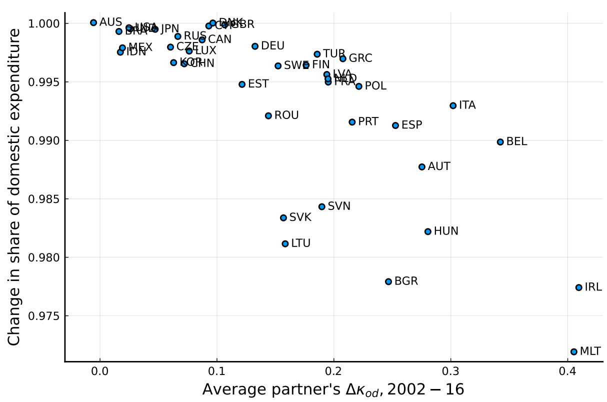
Figure 3: Impact of the change in air connectedness from 2002 to 2016



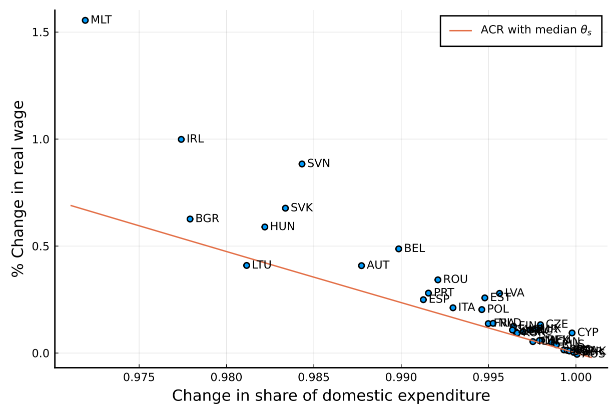
(a) Change in real wages and connectedness



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



(c) Change in domestic expenditure shares and 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 2016. The red solid line in Panel (d) is the welfare changes predicted by the formula $(\widehat{s_{dom}})^{-1/\theta a_L}$, where $\widehat{s_{dom}}$ is the change in overall domestic expenditure share and θ is the median trade elasticity across sectors.

6.3 The Drop in Air Connectedness from 2019 to 2020

The COVID-19 pandemic caused disruptions to international travel. As can be seen from Figure 4, air connectedness with partner countries declined from 2019 to 2020 for all countries in our sample. Among all countries, Ireland and Malta had the largest drop in air connectedness, with an average drop in κ_{od} across partner countries of 0.24 and 0.23, respectively. Another country hit hard by the pandemic was Italy, with a 0.19 decline in air connectedness. We evaluate the impact of such declines using the same approach as in the previous section.

Table 14 reports the results. Consistent with the large declines in κ_{od}, Ireland and Malta saw the largest declines in real wages (0.35% and 0.34%, respectively). Our counterfactual analysis also predicts large declines in real wages in two Eastern European countries, Slovakia and the Czech Republic, which had large (but not the largest) declines in air connectedness and are predicted to have large increases in the share of domestic expenditures.

Figure 4: Average κ_{od} in 2020 v.s. 2019

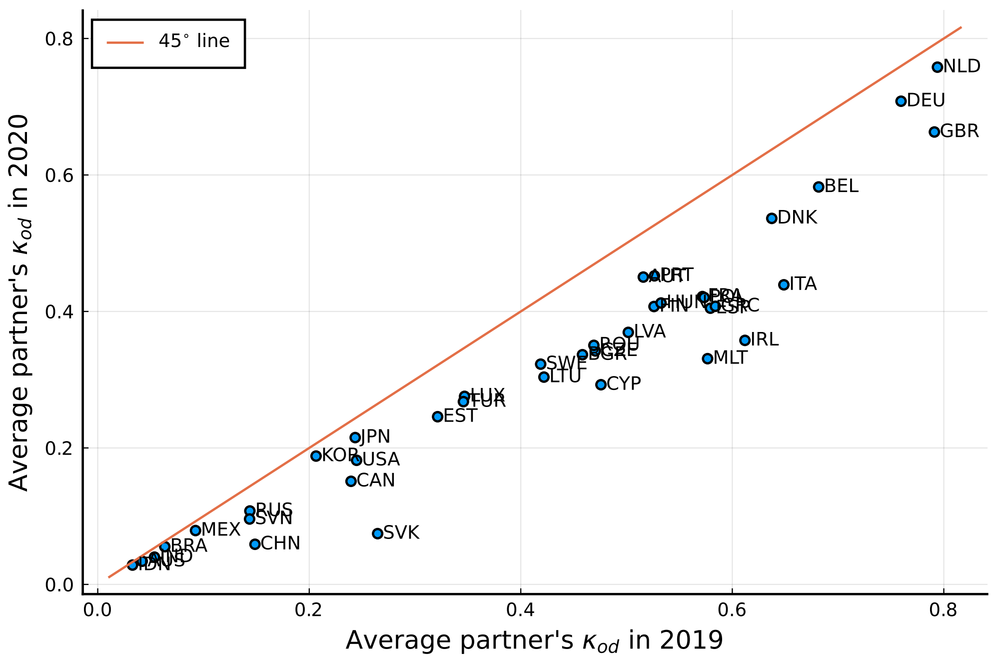


Table 14: Impact of the change in air connectedness from 2019 to 2020

Country	$\kappa_{od,2019}$	$\Delta\kappa_{od}$	$\Delta\%\, s^{final}_{dom}$	$\Delta\%\, s^{input}_{dom}$	$\Delta\%\, s_{dom}$	$\Delta\%\, w/P$
AUS	0.04	-0.01	0.03	0.07	0.05	-0.02
AUT	0.57	-0.06	0.07	0.24	0.14	-0.06
BEL	0.69	-0.10	0.08	0.23	0.15	-0.07
BGR	0.47	-0.11	0.12	0.39	0.25	-0.09
BRA	0.07	-0.01	0.01	0.03	0.02	-0.00
CAN	0.25	-0.08	0.19	0.49	0.32	-0.13
CHN	0.16	-0.09	0.13	0.31	0.25	-0.09
CYP	0.49	-0.18	0.17	0.51	0.27	-0.16
CZE	0.55	-0.13	0.35	0.67	0.53	-0.29
DEU	0.78	-0.04	0.09	0.16	0.12	-0.04
DNK	0.64	-0.09	0.06	0.19	0.12	-0.05
ESP	0.59	-0.16	0.21	0.44	0.32	-0.11
EST	0.33	-0.07	0.11	0.11	0.11	-0.08
FIN	0.54	-0.11	0.13	0.38	0.25	-0.09
FRA	0.61	-0.14	0.23	0.42	0.32	-0.11
GBR	0.80	-0.11	0.19	0.38	0.28	-0.11
GRC	0.60	-0.16	0.11	0.37	0.19	-0.07
HUN	0.55	-0.11	0.21	0.58	0.38	-0.16
IDN	0.04	-0.00	0.01	0.06	0.03	-0.01
IND	0.06	-0.01	0.01	0.03	0.02	-0.00
IRL	0.64	-0.24	0.29	1.37	0.81	-0.35
ITA	0.66	-0.20	0.26	0.41	0.33	-0.12
JPN	0.25	-0.03	0.04	0.04	0.04	-0.01
KOR	0.21	-0.02	0.10	0.21	0.16	-0.05
LTU	0.43	-0.11	0.26	0.40	0.31	-0.11
LUX	0.36	-0.07	0.08	0.41	0.24	-0.14
LVA	0.51	-0.12	0.06	0.08	0.07	-0.05
MEX	0.10	-0.01	0.02	0.10	0.05	-0.01
MLT	0.59	-0.23	0.47	1.33	0.79	-0.34
NLD	0.81	-0.04	0.05	0.17	0.11	-0.03
POL	0.60	-0.14	0.20	0.49	0.34	-0.16
PRT	0.54	-0.07	0.03	0.08	0.05	-0.02
ROU	0.48	-0.11	0.05	0.10	0.07	-0.04
RUS	0.15	-0.04	0.07	0.14	0.10	-0.04
SVK	0.32	-0.21	0.40	0.62	0.52	-0.33
SVN	0.15	-0.04	0.08	0.19	0.13	-0.07
SWE	0.43	-0.09	0.20	0.44	0.31	-0.12
TUR	0.35	-0.07	0.03	0.08	0.05	-0.02
USA	0.25	-0.06	0.04	0.11	0.07	-0.02
Weighted Average	0.36	-0.07				-0.05

Notes: This table reports the impact of the change in air connectedness from 2019 to 2020. $s^{final}_{dom}, s^{input}_{dom}$ 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.

7 Conclusions

This paper shows that communication costs matter in international trade. The increase in direct flights reduces the cost of face-to-face communication, which increases the volume of bilateral trade. We provide both 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 more significant for those products with a low elasticity of substitution, further supporting the mechanism by which direct flights facilitate imports by reducing communication costs, as traded products that are harder to be substituted tend to require more face-to-face interactions to determine the details of the transaction.

We also conduct a counterfactual analysis to calculate the welfare gain from increasing international direct flights and removing air travel regulations. Our model can be used to evaluate other shocks to travel and communication costs, such as the closure of airspace due to the Russia-Ukraine War.

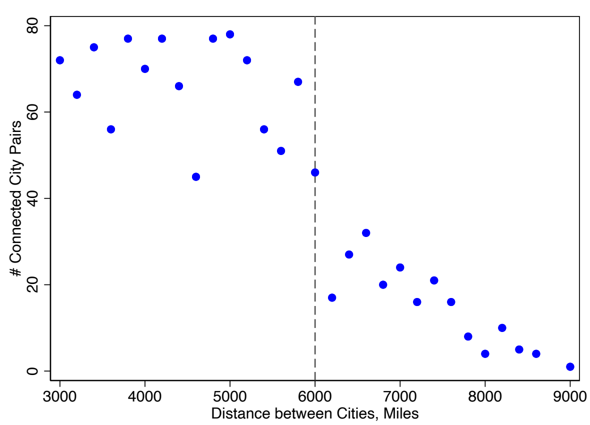
Appendix

A Discontinuity of Connectedness at 6,000 Miles

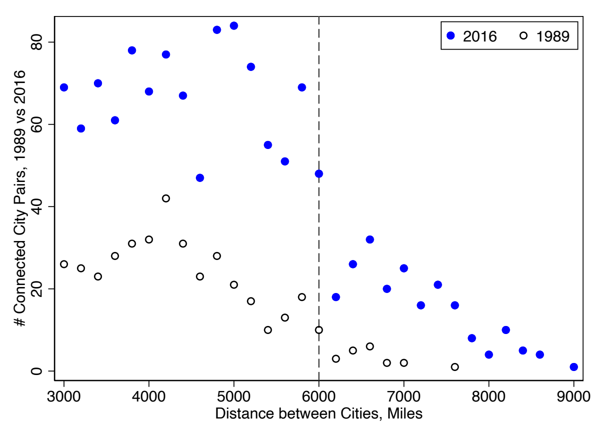
Following Campante and Yanagizawa-Drott (2018), we find that air connectedness between cities drops discontinuously around 6,000 miles due to changes in policy and aviation technology. The United States and Europe passed regulations in the 1950s and 1991, respectively, that restrict hours of operation for flight crew members within a 24-hour period. The regulations lead to additional crew in any flight of more than 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 A-1 confirms the existence of the discontinuity above. Panel A-1a shows a sharp drop in the number of connected city pairs when the distance is larger than 6,000 miles in 2016. Panel A-1b also adds the dots to represent the number of connected city pairs in 1989. We can see the drop is much smoother back to 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.

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



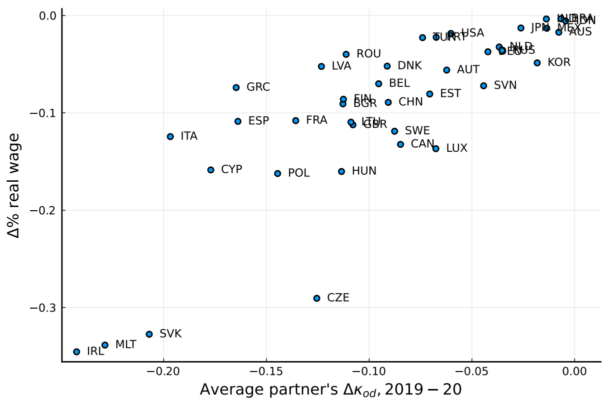
(a) 2016



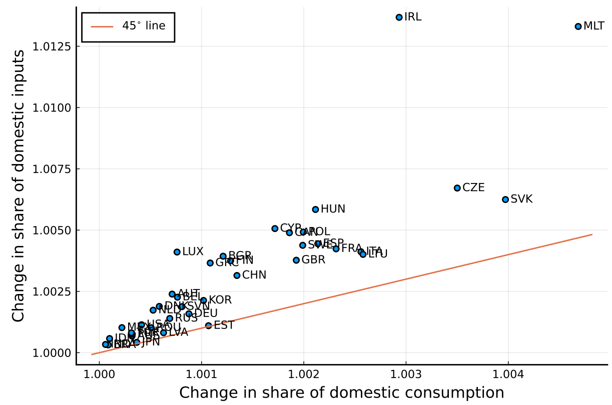
(b) 1989 vs 2016

B Additional Tables and Figures

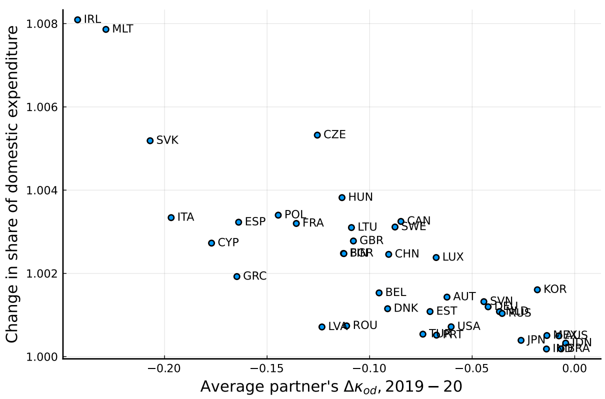
Figure B-1: Impact of the change in air connectedness from 2019 to 2020



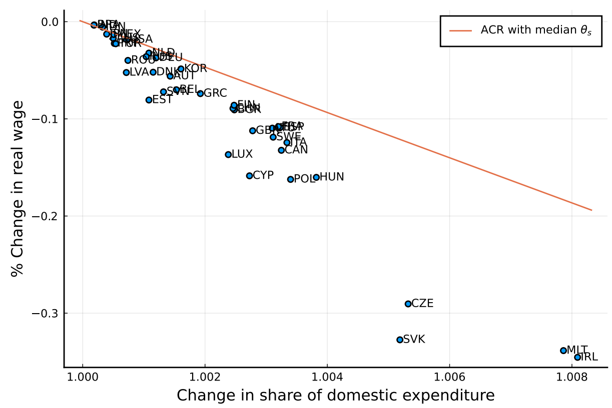
(a) Change in real wages and connectedness



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



(c) Change in domestic expenditure shares and 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 2019 to 2020. The red solid line in Panel (d) is the welfare changes predicted by the formula $(\widehat{s_{dom}})^{-1/\theta a_L}$, where $\widehat{s_{dom}}$ is the change in overall domestic expenditure share and θ is the median trade elasticity across sectors.

References

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.

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.

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.

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

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.

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.

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.

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.

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.

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

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

Nunn, Nathan, "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade," *The Quarterly Journal of Economics*, May 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 2021.

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.